

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: October 23, 2003, 17:09:07 ; Search time 50 Seconds
(without alignments)
3269.767 Million cell updates/sec

Title: US-09-856-681-2

Sequence: 1 MRSEALLYFTLLHFAGAGF.....PPKPSFAPLSTSMKPNDACT 1030

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

- 1: /SIDS1/gcgdata/genseq/genseq_emb1/AA1980. DAT: *
- 2: /SIDS1/gcgdata/genseq/genseq_emb1/AA1981. DAT: *
- 3: /SIDS1/gcgdata/genseq/genseq_emb1/AA1982. DAT: *
- 4: /SIDS1/gcgdata/genseq/genseq_emb1/AA1983. DAT: *
- 5: /SIDS1/gcgdata/genseq/genseq_emb1/AA1984. DAT: *
- 6: /SIDS1/gcgdata/genseq/genseq_emb1/AA1985. DAT: *
- 7: /SIDS1/gcgdata/genseq/genseq_emb1/AA1986. DAT: *
- 8: /SIDS1/gcgdata/genseq/genseq_emb1/AA1987. DAT: *
- 9: /SIDS1/gcgdata/genseq/genseq_emb1/AA1988. DAT: *
- 10: /SIDS1/gcgdata/genseq/genseq_emb1/AA1989. DAT: *
- 11: /SIDS1/gcgdata/genseq/genseq_emb1/AA1990. DAT: *
- 12: /SIDS1/gcgdata/genseq/genseq_emb1/AA1991. DAT: *
- 13: /SIDS1/gcgdata/genseq/genseq_emb1/AA1992. DAT: *
- 14: /SIDS1/gcgdata/genseq/genseq_emb1/AA1993. DAT: *
- 15: /SIDS1/gcgdata/genseq/genseq_emb1/AA1994. DAT: *
- 16: /SIDS1/gcgdata/genseq/genseq_emb1/AA1995. DAT: *
- 17: /SIDS1/gcgdata/genseq/genseq_emb1/AA1996. DAT: *
- 18: /SIDS1/gcgdata/genseq/genseq_emb1/AA1997. DAT: *
- 19: /SIDS1/gcgdata/genseq/genseq_emb1/AA1998. DAT: *
- 20: /SIDS1/gcgdata/genseq/genseq_emb1/AA1999. DAT: *
- 21: /SIDS1/gcgdata/genseq/genseq_emb1/AA2000. DAT: *
- 22: /SIDS1/gcgdata/genseq/genseq_emb1/AA2001. DAT: *
- 23: /SIDS1/gcgdata/genseq/genseq_emb1/AA2002. DAT: *
- 24: /SIDS1/gcgdata/genseq/genseq_emb1/AA2003. DAT: *

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5450	100.0	1030	21	AAV71460	Human semaphorin 6
2	5103.5	93.6	975	22	AAB90731	Human C14orf1 protein
3	5088	93.4	974	19	AAW64221	Human secreted protein
4	4921	90.3	939	21	AAW62030	Human semaphorin 1
5	4590.5	84.2	884	21	AAW20511	Human semaphorin 1
6	3677	67.5	699	22	AAW95139	Human semaphorin 1
7	3373	61.9	530	21	AAW20443	Human semaphorin 1
8	2980.5	54.7	574	22	AAW95217	Human protein sequence
9	2952	54.2	562	22	AAW94104	Human protein sequence

44	1386.5	25.4	367	22	ABG93769	Rat semaphorin Y.
43	1386.5	25.4	367	22	ABG93769	Rat semaphorin Y.
42	1519.5	27.9	517	24	AAW57280	Human semaphorin Y
41	1567.5	28.8	517	24	ABU03537	Angiogenesis-associ
40	1567.5	28.8	517	22	AAE85338	Human LF221 secret
39	1611	29.6	476	23	AAE83961	Human semaphorin R
38	1611	29.6	476	23	ABG79172	Human semaphorin-L
37	1611	30.3	321	20	AAW22114	Mouse ZSMF-3 prote
36	1714	31.4	321	20	AAW22113	Human ZSMF-3 prote
35	1857	34.1	686	22	AAE85339	Human semaphorin R
34	1923	35.3	641	22	AAE63215	Partial sequence o
33	1989.5	36.5	712	23	ABG79174	Human semaphorin-L
32	1997	36.6	887	18	AAW19856	Rat semaphorin Z.
31	2051.5	36.7	888	24	ABU03558	Angiogenesis-associ
30	2008.5	36.9	768	24	ABG79176	Human semaphorin-L
29	2008.5	36.9	768	24	ABU59922	Novel secreted ant
28	2009.5	36.9	888	24	ABU67117	Human secreted/titr
27	2009.5	36.9	888	24	ABU67841	Human PRO polyepi
26	2009.5	36.9	888	24	ABU69110	Human PRO polyepi
25	2009.5	36.9	888	24	ABU59037	Human PRO polyepi
24	2009.5	36.9	888	23	ABBA4350	Human PRO433 prot
23	2009.5	36.9	888	23	ABBA9556	Human angiogenesis
22	2009.5	36.9	888	22	AAU12443	Human PRO433 poly
21	2009.5	36.9	888	18	AAW19857	Human semaphorin Z
20	2013.5	40.6	416	22	AAE85886	Human protein seg
19	2213	42.7	1033	23	ABG79173	Human semaphorin 6
18	2326.5	42.8	1035	23	ABG79175	Human semaphorin 6
17	2335	42.8	1017	23	AAE79413	CADHP-2, Incyte II
16	2346	43.1	1088	23	ABG79177	Human semaphorin-L
15	2349.5	43.3	1086	22	AAE63213	Amino acid sequen
14	2621.5	48.1	507	22	AAE82668	Human protein seg
13	2679.5	52.8	962	22	ABG04670	Novel human diagn
12	2879.5	53.0	863	22	ABG04664	Novel human diagn
11	2891	54.0	562	22	AAW33444	Human polypeptide,
10	2947	54.4	562	22	AAW33446	Human polypeptide,

ALIGNMENTS

XX	RESULT 1
XX	AAV71460
XX	ID AAV71460 standard; Protein; 1030 AA.
XX	AAV71460;
XX	04-OCT-2000 (first entry)
XX	Human semaphorin 6A-1.
XX	Human; semaphorin 6A-1; (HSA) SEMA6A-1; neuronal development; apoptosis
XX	neuronal regeneration; Ena/VASP protein family; immunomodulatory;
XX	gene therapy; diagnostic agent; therapeutic agent; differentiation;
XX	cytoskeletal stabilization; plasticity.
XX	Homo sapiens.
XX	
XX	Key
XX	Binding-site
XX	959..1030
XX	/note= "Zyxin-like domain that selectively binds to
XX	members of Ena/VASP protein family, especially Evi1"
XX	957..961
XX	/note= "Specific binding motif for members of
XX	Ena/VASP protein family, especially Evi"
XX	1009..1014
XX	/note= "Specific binding motif for members of
XX	Ena/VASP protein family, especially Evi1"
XX	
XX	W0200031252-A1.
XX	
XX	02-JUN-2000.
XX	

e.g. cytokine and cell proliferation/differentiation activity, the immune system and hematopoiesis regulating activity -

Disclosure: Page 487-490; 557pp; English.

Human cDNA clones represented in AAF98374 - AAF98489 encode secreted proteins AAB90667 - AAB90750. The cDNA clones are isolated from various tissue types, and may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate protein expression. The polypeptides and nucleic acids may be used as nutrients or to modulate cytokine and cell proliferation/differentiation activity and may also be involved in modulation of the immune system. The cDNA sequences encode proteins, their agonists and/or antagonists exhibit haematopoiesis regulating activity; tissue growth activity; activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic and thrombolytic activity; receptor/ligand activity; anti-inflammatory activity; and/or haematopoiesis activity; cadherin/tumour suppressor activity; and/or tumour inhibition activity. Included in the invention are probes represented in AAF98490 - AAF98572 which are specific for the cDNA clones encoding the secreted proteins.

Sequence 975 AA:

Query Match 93.6%; Score 5103.5; DB 22; Length 975;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 972; Conservative 0; Mismatches 3; Indels 55; Gaps 1;

1 MREBALLVFTLLHFGAGPDESEPTISGNYTKQYFVGHKQGRNTTQRRLDIOM 60
1 MREBALLVFTLLHFGAGPDESEPTISGNYTKQYFVGHKQGRNTTQRRLDIOM 60
61 IMINGTLYIARADHITVDITDTSHEIYCSKKLTWKSQADVDTCRMKRIKDCNHF 120
61 IMINGTLYIARADHITVDITDTSHEIYCSKKLTWKSQADVDTCRMKRIKDCNHF 120
121 IKVLLKNDALFYCGTNAFPCSNRYKMDLFEFGDFSGMARCPYDAHANVALPADG 180
121 IKVLLKNDALFYCGTNAFPCSNRYKMDLFEFGDFSGMARCPYDAHANVALPADG 180
181 KLYSATYDPLADAVIYRSIGESPTLRVYKDSKMLKEPFYQAVYGDYIYFFREIA 240
181 KLYSATYDPLADAVIYRSIGESPTLRVYKDSKMLKEPFYQAVYGDYIYFFREIA 240
241 VEYNTMGKVPFPRVAQVCKNDMGSGORVLEKQWTSFLKARLNGSVPGDSHFENILQAVT 300
241 VEYNTMGKVPFPRVAQVCKNDMGSGORVLEKQWTSFLKARLNGSVPGDSHFENILQAVT 300
301 DVIRINGADVLAFTSTPYNSIPGSAVCAVMDLIASVFTGRFKEQKSPDSTWTPVDER 360
301 DVIRINGADVLAFTSTPYNSIPGSAVCAVMDLIASVFTGRFKEQKSPDSTWTPVDER 360
361 VPKPRPGCCAGSSSLERYATSNRPDDTLNFKTHPLMDEAVPSIENRPNELRTWRYRL 420
361 VPKPRPGCCAGSSSLERYATSNRPDDTLNFKTHPLMDEAVPSIENRPNELRTWRYRL 420
421 TKIADTAAGPYQNTVTVFLGSEKGIILKFLARIGNSGFLNDSFLFEMSVNSKCYD 480
421 TKIADTAAGPYQNTVTVFLGSEKGIILKFLARIGNSGFLNDSFLFEMSVNSKCYD 480
481 GVEDKRIMGMQJDRASSSLIYVAFSTCVTLNPLGRCEHCKCKCTCIASDPYCGMIKEGG 540
481 GVEDKRIMGMQJDRASSSLIYVAFSTCVTLNPLGRCEHCKCKCTCIASDPYCGMIKEGG 540
541 ACSHSPNSRLTFEODIERGNTDGLGDCNSFVALNGHSSSLPFTTSDTAQEGYESR 600
541 ACSHSPNSRLTFEODIERGNTDGLGDCNSFVALNGHSSSLPFTTSDTAQEGYESR 600
601 GGMLDKHLIDSPDSTPLGAVSNHNDKQKGVIRSYLKGHQDQVPTLLAIVTLAV 660
601 GGMLDKHLIDSPDSTPLGAVSNHNDKQKGVIRSYLKGHQDQVPTLLAIVTLAV 660
661 MGAVSGITVYCVCHRRKDVAVVQKKEKELTSSRGSMSSVTKSLGDFDTQSKDPKE 720
661 MGAVSGITVYCVCHRRKDVAVVQKKEKELTSSRGSMSSVTKSLGDFDTQSKDPKE 720

606 MGAVSGITVYCVCHRRKDVAVVQKKEKELTSSRGSMSSVTKSLGDFDTQSKDPKE 665
721 ALTPLEHNGKATPGTAXKMLKADQHHDLTALPTPESTPTLQQRKPSRSGREWRN 780
666 ALTPLEHNGKATPGTAXKMLKADQHHDLTALPTPESTPTLQQRKPSRSGREWRN 725
781 QNLINACTKMPGSPVITPDLPLRASPSHIBSVVLPITQOQYQHEVVOPKMSEYAO 840
726 QNLINACTKMPGSPVITPDLPLRASPSHIBSVVLPITQOQYQHEVVOPKMSEYAO 785
841 MALEDQAATLEKTKIKELLSKSPNHGYNLVEINDSLPPKYQPRASIGPPASISQTL 900
786 MALEDQAATLEKTKIKELLSKSPNHGYNLVEINDSLPPKYQPRASIGPPASISQTL 845
901 SKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLSRNOSFGRGDNPP 960
846 SKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLSRNOSFGRGDNPP 905
961 PAPQFVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKTPTLKPVPKPSFAPS 1020
906 PAPQFVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKTPTLKPVPKPSFAPS 965
1021 TSMKPNDACT 1030
966 TSMKPNDACT 975

RESULT 3

AAW64221
ID AAW64221 standard; protein; 974 AA.

AAW64221;

06-OCT-1998 (first entry)

Human secreted protein from clone CUI45_1.

Secreted protein; human fetal brain; nutrition; cytokine; stimulant;
cell proliferation; differentiation; immune system; suppressor; ligand;
regulator; hematopoiesis; tissue growth; activin; inhibin; haemostatic;
chemotaxis; chemokinetic; thrombosis; receptor; cadherin; tumour;
anti-inflammatory.

Homo sapiens.

MO9827205-A2.

25-JUN-1998.

17-DEC-1997; 97WO-US23330.

16-DEC-1997; 97US-0991872.

18-DEC-1996; 96US-0769192.

13-JAN-1997; 97US-0783401.

(GENEY) GENETICS INST INC.

Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
Racie LA, Spaulding V, Treacy M;

WPI; 1998-362774/31.

N-PSDB; AAV44295.

New polynucleotides and secreted proteins - obtained from human
foetal brain, human adult testes, human adult brain and human adult
salivary gland cDNA libraries

Claim 17; Page 71-74; 110pp; English.

This sequence represents a novel secreted protein from clone CUI45_1
isolated from a human fetal brain cDNA library. This protein has
applications for nutritional use, cytokine and cell
proliferation/differentiation activity, immune stimulating or

suppressing activity, hematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombotic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, tumour inhibition activity and other activities.

XX Sequence 974 AA;

Query Match 93.4%; Score 5088; DB 19; Length 974;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 974; Conservative 0; Mismatches 3; Indels 56; Gaps 2;

1 MRSBALLYFTLLHFAAGPEDESEPISSHNATKQYVVFVGHKPGNTTQRRLDIOM 60
1 MRSBALLYFTLLHFAAGPEDESEPISSHNATKQYVVFVGHKPGNTTQRRLDIOM 60
61 IMVNGTLYAARHITVVDIDTSHTEETYSKLTWKSQADVDTCRMKGKDECHNF 120
61 IMVNGTLYAARHITVVDIDTSHTEETYSKLTWKSQADVDTCRMKGKDECHNF 119
61 IMVNGTLYAARHITVVDIDTSHTEETYSKLTWKSQADVDTCRMKGKDECHNF 119
121 IKVLLKKNDDALFVCGTNAFNPSCNRYKMDLPEFGDEFGQMARCPYDAKANVALPADG 180
120 IKVLLKKNDDALFVCGTNAFNPSCNRYKMDLPEFGDEFGQMARCPYDAKANVALPADG 179
181 KLYSATVVDPLAIDAVIYRSLGESPTLTVKHDSKMLKEPIFVQAVDYGVITYFFFRFA 240
180 KLYSATVVDPLAIDAVIYRSLGESPTLTVKHDSKMLKEPIFVQAVDYGVITYFFFRFA 239
241 VEYNTMGKVPFRRVQVQCNKMGSORLEKQWTSFLKARLNCSPGDSHFENLQAVT 300
240 VEYNTMGKVPFRRVQVQCNKMGSORLEKQWTSFLKARLNCSPGDSHFENLQAVT 299
301 DVIRINGRDVATSTFPYNSIPGSAVAYMDLADIASVFTGRFKQKSPDSTWTPVDER 360
300 DVIRINGRDVATSTFPYNSIPGSAVAYMDLADIASVFTGRFKQKSPDSTWTPVDER 359
360 VPKPRPGCCAGSSSERVATSNPEPDDTLNFKTHPLMDEAVPSIFNRWFLRTWRYRL 419
421 TKIADVTAGPQNTHTVVLGSEKGIILKFLARIGSGFLNLSLFLSEMSVNSEKSYD 480
420 TKIADVTAGPQNTHTVVLGSEKGIILKFLARIGSGFLNLSLFLSEMSVNSEKSYD 479
481 GVEDKRMGQDRASSLYVAFSTCVIYVPLGRCHGCKCTCIASSDPYCGWIKKCG 540
480 GVEDKRMGQDRASSLYVAFSTCVIYVPLGRCHGCKCTCIASSDPYCGWIKKCG 539
541 ACSHLSPNRSLTEQDIERGNTDGLGDCNHSFVALN----- 575
540 ACSHLSPNRSLTEQDIERGNTDGLGDCNHSFVALN----- 575
601 GGMIDMKHLDPDSIDPLGAVSSHNQDKGVIRESYLKGHDQVPTLLAIAVILAVY 660
576 -----GVIRBSYLKGHDQVPTLLAIAVILAVY 604
661 MGAVFSGITVYCVCDHRKDVAVVORKEKELTHSRGSSSVYTKLSGLFSDTQSKPKPE 720
605 MGAVFSGITVYCVCDHRKDVAVVORKEKELTHSRGSSSVYTKLSGLFSDTQSKPKPE 664
721 AILTPMHNGKLTAPNTAKMLIADQHHDLTLPTPESTPTLQQRKRSRGRKERN 780
665 AILTPMHNGKLTAPNTAKMLIADQHHDLTLPTPESTPTLQQRKRSRGRKERN 724
781 QNLINACTKMPNMGSPVITPDLPLRASPSHISPVVLLPTIQGYQHEVYVQPMSEVAQ 840
725 QNLINACTKMPNMGSPVITPDLPLRASPSHISPVVLLPTIQGYQHEVYVQPMSEVAQ 784
841 MALDQAAITLEYKTIKEHLSKSPNMGVNTVENLDSPPKYVQGEASIGPPGASLSOTGL 900
785 MALDQAAITLEYKTIKEHLSKSPNMGVNTVENLDSPPKYVQGEASIGPPGASLSOTGL 844
901 SKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKKNNTNSNSHLSRNGSFGRGNDP 960

DB 845 SKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKKNNTNSNSHLSRNGSFGRGNDP 904
QY 961 PAPQVDSIQVWSSSQSGQAVVRSOPSLNAYNSLTRSGLKSTPSLKPVPKPSFAPLS 1020
DB 905 PAPQVDSIQVWSSSQSGQAVVRSOPSLNAYNSLTRSGLKSTPSLKPVPKPSFAPLS 964
QY 1021 TSMKENDACT 1030
DB 965 TSMKENDACT 974

RESULT 4

ID AAB23030 standard; Protein: 939 AA.

AC AAB23030;

DT 16-JAN-2001 (first entry)

DB Human semaphorin protein-like splice variant, SECX 2864933-1.

XX SECX protein; human; secreted; membrane-associated; cancer;
XX proliferation regulator; differentiation regulator; non-malignant tumour;
XX immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;
XX infection; inflammatory disorder; arthritis; haematopoietic disorder;
XX skin disorder; cardiovascular disorder; atherosclerosis; restenosis;
XX neurological disease; Alzheimer's disease; trauma; wounding;
XX spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
XX anti-HIV; anti-inflammatory; antidiabetic; antiarteriosclerotic;
XX neuroprotective; vulnereary; antiallergic; antimicrobial; cardiant;
XX dermatological; gene therapy.

OS Homo sapiens.

PN WO20053742-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06280.

XX 09-MAR-1999; 99US-0123667.

XX 08-MAR-2000; 2000US-0123667.

PA (CURA-) CURAGEN CORP.

PI Shinkens RA;

DR WPI; 2000-594318/56.

DR N-PSDB; AAA93617.

PT Novel human membrane associated or secreted polypeptides and
PT polynucleotides useful for diagnosis, prevention and treatment of
PT pathological states such as cancer, immune, cardiovascular and
PT neurological disorders

PS Claim 1; Fig 2; 151pp; English.

XX Sequences AAB23029-B23048 represent human SECX proteins. The SECX
XX proteins and act as regulator of cellular proliferation and
XX differentiation. SECX proteins or nucleotides are useful for diagnosing
XX the presence of, or predisposition to, a disease associated with altered
XX levels of SECX proteins and nucleotides. The SECX proteins are also
XX useful to screen compounds that modulate SECX activity or expression. The
XX interaction of a SECX protein with other cellular proteins may be useful
XX to modulate the activity of a partner protein. SECX nucleotides are useful
XX for the recombinant expression and cell survival. SECX nucleotides are useful
XX for the recombinant expression of SECX protein, and may be used to detect
XX SECX mRNA or genetic lesions in the SECX gene. They may also be used to
XX modulate SECX expression (e.g., using antisense oligonucleotides). SECX
XX nucleic acid sequences are also useful for identifying a cell or tissue
XX type in a biological sample, and in forensic biology. SECX primers or

CC probes are useful for detecting the presence of SECC nucleotides and for
 CC screening tissue cultures for contamination. Diseases that may be treated
 CC or prevented using SECC proteins or nucleotides include cancer (e.g.,
 CC colorectal carcinoma, prostate cancer), benign tumours, immune disorders
 CC (including autoimmune diseases, transplant rejection, allergies, AIDS),
 CC infections, inflammatory disorders, arthritis, haematopoietic disorders,
 CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
 CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
 CC surgical or traumatic wounds, spinal cord injury), and skeletal
 CC disorders.

XX Sequence 939 AA;

Query Match 90.3%; Score 4921; DB 21; Length 939;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 928; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MRSEALLLYTLHFAAGAFPPDSEPISSHGNTKQYVFGHKGGRNTTORHLDIQM 60
DB 1 MRSEALLLYTLHFAAGAFPPDSEPISSHGNTKQYVFGHKGGRNTTORHLDIQM 60
QY 61 IMNMGTLIARADHIYVDIDTSHTEETIYCSKLTWKSQADVDTCMKKGKHECHNF 120
DB 61 IMNMGTLIARADHIYVDIDTSHTEETIYCSKLTWKSQADVDTCMKKGKHECHNF 120
QY 121 IKVLLKQNDALFVCGTNAFNPSCRNYKMDLEPFDEFSGNARCPYDAKANAVALFADG 180
DB 121 IKVLLKQNDALFVCGTNAFNPSCRNYKMDLEPFDEFSGNARCPYDAKANAVALFADG 180
QY 181 KLYSAVTVDFLAIDAVYRSLSGSPPLRTVKHDSKWLKEPYQVADYDYYFFPRETA 240
DB 181 KLYSAVTVDFLAIDAVYRSLSGSPPLRTVKHDSKWLKEPYQVADYDYYFFPRETA 240
QY 241 VERNMGKVPFPAVAVOCNMDMGSOBYLEKQWTSFLKRLNCSVPGSDHFEYNILQAVT 300
DB 241 VERNMGKVPFPAVAVOCNMDMGSOBYLEKQWTSFLKRLNCSVPGSDHFEYNILQAVT 300
QY 301 DVIRINGRDVLAFTSTPYNSIPGSAVCAVMDLIASVFTGRFKEKSPDSTWTPVDER 360
DB 301 DVIRINGRDVLAFTSTPYNSIPGSAVCAVMDLIASVFTGRFKEKSPDSTWTPVDER 360
QY 361 VPPRRGCCGSSSLRYATSNFPDDTNFIKTHLMDEAVSIFNRPFELTWRYRL 420
DB 361 VPPRRGCCGSSSLRYATSNFPDDTNFIKTHLMDEAVSIFNRPFELTWRYRL 420
QY 421 TKTAVIDTAAGPYQNHVYFLGSEKGIILKFLARIGNSGFLNDSLPLEMSVNSSEKSYD 480
DB 421 TKTAVIDTAAGPYQNHVYFLGSEKGIILKFLARIGNSGFLNDSLPLEMSVNSSEKSYD 480
QY 481 GVEDKRIWQDLRASSSLVAFSTCVIKVPLGRCERHCKCKTCTASRPYCGMIYEGG 540
DB 481 GVEDKRIWQDLRASSSLVAFSTCVIKVPLGRCERHCKCKTCTASRPYCGMIYEGG 540
QY 541 ACSHLSPNSRLTEODIERGNTDGLGCHNSFVALNHSLSLPTSTTSSTQEGYESR 600
DB 541 ACSHLSPNSRLTEODIERGNTDGLGCHNSFVALNHSLSLPTSTTSSTQEGYESR 600
QY 601 GGMMDWCHLLDSPSTPLGAVSSHNQDKGVIRESYLKGHDQVLEVTLLAVALIAVAFV 660
DB 601 GGMMDWCHLLDSPSTPLGAVSSHNQDKGVIRESYLKGHDQVLEVTLLAVALIAVAFV 660
QY 661 MGAVFSGITVYCVCDHRKDVAVVQRKEKELTHSRSGSSSVTKSLGFGSTQKDPPE 720
DB 661 MGAVFSGITVYCVCDHRKDVAVVQRKEKELTHSRSGSSSVTKSLGFGSTQKDPPE 720
QY 721 AILPLIMNGKLAIPGNTAKMLIKADQHLHLTLPLPESTPTLQOKKRPGRGSEWERN 780
DB 721 AILPLIMNGKLAIPGNTAKMLIKADQHLHLTLPLPESTPTLQOKKRPGRGSEWERN 780
QY 781 ONLINACTKQMPKPSVPIPTDLPLRASPSHIPSVVVLPIQOGYQHEVYDQPKMSEVAQ 840
DB 781 ONLINACTKQMPKPSVPIPTDLPLRASPSHIPSVVVLPIQOGYQHEVYDQPKMSEVAQ 840

```

```

QY 841 MALEDQATLEYKTIKEHLSSKSPNHNVLNENIDSLPPKVPQREASIGPPGASLSQTGL 900
DB 841 MALEDQATLEYKTIKEHLSSKSPNHNVLNENIDSLPPKVPQREASIGPPGASLSQTGL 900
QY 901 SKRLEMHSSSYGVYKRSYPTNSLTSSHQAT 932
DB 901 SKRLEMHSSSYGVYKRSYPTNSLTSSHQAT 932

```

RESULT 5

AAAB23031 ID AAAB23031 standard; Protein, 884 AA.

XX AC AAAB23031;

XX DT 16-JAN-2001 (first entry)

XX DE

Human semaphorin protein-like splice variant, SECC 2864933-2.

SECC protein; human; secreted; membrane-associated; cancer;
 proliferation regulator; differentiation regulator; non-malignant tumour;
 immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;
 infection; inflammatory disorder; arthritis; haematopoietic disorder;
 skin disorder; cardiovascular disorder; atherosclerosis; restenosis;
 neurological disease; Alzheimer's disease; trauma; wounding;
 spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
 anti-HIV; antiinflammatory; antiarthritic; antiatheriosclerotic;
 neuroprotective; vulnerability; antiallergic; antimicrobial; cardiant;
 dermatological; gene therapy.

XX OS Homo sapiens.

XX PN WO20005742-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US06280.

XX ER 09-MAR-1999; 99US-0123667.
 XX PR 08-MAR-2000; 2000US-0123667.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkete RA;

XX DR WPI, 2000-594318/56.

XX DR N-PSDB; AAA93618.

PT Novel human membrane associated or secreted polypeptides and
 PT polynucleotides useful for diagnosis, prevention and treatment of
 PT pathological states such as cancer, immune, cardiovascular and
 PT neurological disorders -

XX Claim 1; Fig 3; 15pp; English.

XX Sequences AAAB23029-B23048 represent human SECC proteins. The SECC
 CC proteins and act as regulator of cellular proliferation and
 CC differentiation. SECC proteins or nucleotides are useful for diagnosing
 CC the presence of, or predisposition to, a disease associated with altered
 CC levels of SECC proteins and nucleotides. The SECC proteins are also
 CC useful to screen compounds that modulate SECC activity or expression. The
 CC interaction of a SECC protein with other cellular proteins may be useful
 CC to modulate the activity of a partner protein. SECC nucleotides are useful
 CC for the recombinant expression of SECC protein, and may be used to detect
 CC SECC mRNA or genetic lesions in the SECC gene. They may also be used to
 CC modulate SECC expression (e.g., using antisense oligonucleotides). SECC
 CC nucleic acid sequences are also useful for identifying a cell or tissue
 CC type in a biological sample, and in forensic biology. SECC primers or
 CC probes are useful for detecting the presence of SECC nucleotides and for
 CC screening tissue cultures for contamination. Diseases that may be treated
 CC or prevented using SECC proteins or nucleotides include cancer (e.g.,

CC colorectal carcinoma, prostate cancer), benign tumours, immune disorders
 CC (including autoimmune diseases, transplant rejection, allergies, AIDS),
 CC infections, inflammatory disorders, arthritis, haematopoietic disorders,
 CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
 CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
 CC surgical or traumatic wounds, spinal cord injury), and skeletal
 CC disorders.

XX Sequence 884 AA;

Query Match 84.2%; Score 4590.5; DB 21; Length 884;

Best Local Similarity 93.7%; Pred. No. 0; Mismatches 2; Indels 55; Gaps 1;

Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;

QY 1 MRSBALLYFTLHPAGAGPEDESEPISSHGNTKQYVFGHKPGRNTQRRLDIOM 60
 Db 1 MRSBALLYFTLHPAGAGPEDESEPISSHGNTKQYVFGHKPGRNTQRRLDIOM 60
 QY 61 IMVNGTLYIARDHIYTVDDIDTSHTEBIYCSKLTWKSQADVDTCMKKXKDECHNF 120
 Db 61 IMVNGTLYIARDHIYTVDDIDTSHTEBIYCSKLTWKSQADVDTCMKKXKDECHNF 120
 QY 121 IKVLKKNDDALFVCGNNAENPCSNYKMDLPEPGDFSGMARCPYDAKIANALPADG 180
 Db 121 IKVLKKNDDALFVCGNNAENPCSNYKMDLPEPGDFSGMARCPYDAKIANALPADG 180
 QY 181 KLYSATVDFLAIDAVIYRSLSGESPTLRTVXHSKWLKEPYFVQAVDGYDIYFFPREIA 240
 Db 181 KLYSATVDFLAIDAVIYRSLSGESPTLRTVXHSKWLKEPYFVQAVDGYDIYFFPREIA 240
 QY 241 VEYNTGKVPPRAQVCKNMGSQRVLEKQWTSFLKARLNCSPGSHHYFNILQAVT 300
 Db 241 VEYNTGKVPPRAQVCKNMGSQRVLEKQWTSFLKARLNCSPGSHHYFNILQAVT 300
 QY 301 DVIRNGRDVLTATSTPYNSIPGSAVCAVMDLJASVFTGFRKQKSPDSTTTPVPPER 360
 Db 301 DVIRNGRDVLTATSTPYNSIPGSAVCAVMDLJASVFTGFRKQKSPDSTTTPVPPER 360
 QY 361 VKPRPGCCAGSSSLERYATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWELRWRYRL 420
 Db 361 VKPRPGCCAGSSSLERYATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWELRWRYRL 420
 QY 421 TKIADVTAAGPYQNTHTVFLGSEKGIILKFLARIGNSGLNLSLLEMSVYNSKSCYD 480
 Db 421 TKIADVTAAGPYQNTHTVFLGSEKGIILKFLARIGNSGLNLSLLEMSVYNSKSCYD 480
 QY 481 GVEDRIMOMOLDRASSSLIYVAFSTCVIKVPLGRGERHCKCKTICIASRDPYCGWIKESG 540
 Db 481 GVEDRIMOMOLDRASSSLIYVAFSTCVIKVPLGRGERHCKCKTICIASRDPYCGWIKESG 540
 QY 541 ACSHTSPNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTTSOSTAOGYESR 600
 Db 541 ACSHTSPNSRLTFEODIERGNTDGLDCHNSFVALNGHSSLLPSTTTSOSTAOGYESR 600
 QY 601 GGNLDMKHLIDBPDSTDLPLGAVSSHKQKQKVIKRSYKKGHDQVPTVLLAIVLAVFV 660
 Db 601 GGNLDMKHLIDBPDSTDLPLGAVSSHKQKQKVIKRSYKKGHDQVPTVLLAIVLAVFV 660
 QY 661 MGAVSGITVYCVCDHRKDVAVVORKEKELTSSRGSMSVTKLSGLFGDTQSKDPKE 720
 Db 661 MGAVSGITVYCVCDHRKDVAVVORKEKELTSSRGSMSVTKLSGLFGDTQSKDPKE 720
 QY 721 AITPLMNGKATPNTAMMLIKADQHLIDTALPTPESTPTLQOKRPSRGRSWERN 780
 Db 721 AITPLMNGKATPNTAMMLIKADQHLIDTALPTPESTPTLQOKRPSRGRSWERN 780
 QY 781 QNLIINACTKMPMGSPVPTDPLASPSHISVVVLLPTQGYGHEVYDOPKMEVQA 840
 Db 781 QNLIINACTKMPMGSPVPTDPLASPSHISVVVLLPTQGYGHEVYDOPKMEVQA 840
 QY 841 MALEDQAATLEYKTIKELSSKSPNGVNLVENLDSLPKVPVPREASLGPAGSLSGTGI 900
 Db 841 MALEDQAATLEYKTIKELSSKSPNGVNLVENLDSLPKVPVPREASLGPAGSLSGTGI 900
 QY 900 MALEDQAATLEYKTIKELSSKSPNGVNLVENLDSLPKVPVPREASLGPAGSLSGTGI 900
 Db 900 MALEDQAATLEYKTIKELSSKSPNGVNLVENLDSLPKVPVPREASLGPAGSLSGTGI 900

QY 901 SKRLEMHSSSYGVDRKSYPTNLSLRSHQAT 932
 Db 846 SKRLEMHSSSYGVDRKSYPTNLSLRSHQAT 877

RESULT 6

AAB95139 standard; Protein; 699 AA.

AAB95139; 26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:17154.

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000CEP-0116126.

29-JUN-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-018776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602

full-length cDNAs defined in the specification, and for the detection

and/or diagnosis of the abnormality of the proteins encoded by the

full-length cDNAs -

Claim 8; SEQ ID 17154; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification, where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary

to the complementary strand of a polynucleotide which comprises one of

the 5602 nucleotide sequences defined in the specification, where the

oligonucleotide comprises at least 15 nucleotides; or (b) a combination

of an oligonucleotide comprising a sequence complementary to the

complementary strand of a polynucleotide which comprises a 5'-end

sequence and an oligonucleotide comprising a sequence complementary to a

polynucleotide which comprises a 3'-end sequence, where the

oligonucleotide comprises at least 15 nucleotides; and the combination

of the 5'-end sequence/3'-end sequence is selected from those defined in

the specification. The primer sets can be used in antisense therapy and

in gene therapy. The primers are useful for synthesizing polynucleotides,

particularly full-length cDNAs. The primers are also useful for the

detection and/or diagnosis of the abnormality of the proteins encoded by

the full-length cDNAs. The primers allow obtaining of the full-length

cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

AAB95893 represent human amino acid sequences; and AAH13629 to AAH1632

represent oligonucleotides, all of which are used in the exemplification

of the present invention.

Sequence 699 AA;

Query Match 67.5%; Score 3677; DB 22; Length 699;
 Best Local Similarity 100.0%; Pred. No. 1.6e-310;

Matches	699;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	332	MLDIASVFTGPEKESKSPDSTWTPVDEVPKPRPCCCAGSSSLERATSNFPPDTLNF	391						
Db	1	MLDIASVFTGPEKESKSPDSTWTPVDEVPKPRPCCCAGSSSLERATSNFPPDTLNF	60						
QY	392	ITHTPLMDAVALSINRPFELRTMRYRLTKIAVDTAAGPYQNHVYFELSGKJIIKFL	451						
Db	61	IXTHPLMDAVALSINRPFELRTMRYRLTKIAVDTAAGPYQNHVYFELSGKJIIKFL	120						
QY	452	ATIGNSGFINDSLFLEEMSVYNSEKCSYDVEDKRLMGQDLRASSSLVYASTCIVKP	511						
Db	121	ATIGNSGFINDSLFLEEMSVYNSEKCSYDVEDKRLMGQDLRASSSLVYASTCIVKP	180						
QY	512	LGRCEHSGCKTKCIASRDYCGWIKGACSHLSNPSRLTEODIERGNTGLDCHNS	571						
Db	181	LGRCEHSGCKTKCIASRDYCGWIKGACSHLSNPSRLTEODIERGNTGLDCHNS	240						
QY	572	FVALNGHSSGLPSTTSSSTAOEGESRGMLDMGHLDPDSTDPGLAVSSHNDKK	631						
Db	241	FVALNGHSSGLPSTTSSSTAOEGESRGMLDMGHLDPDSTDPGLAVSSHNDKK	300						
QY	632	GVIRESYLKGHDQVAVTLLAIVLLAFVWGAVERSGITVYCVCDHRKQVAVVQREKEL	691						
Db	301	GVIRESYLKGHDQVAVTLLAIVLLAFVWGAVERSGITVYCVCDHRKQVAVVQREKEL	360						
QY	692	THSRGSMSSVTLKSLFGDTQSKDPKPEAILPLMHNGKLATPNTAKYLIRADQHHLD	751						
Db	361	THSRGSMSSVTLKSLFGDTQSKDPKPEAILPLMHNGKLATPNTAKYLIRADQHHLD	420						
QY	752	L7ALPPESTPTLQQRKPKRSGREMERONLINACTCKMPKPPSPITPDLPLRASPSH	811						
Db	421	L7ALPPESTPTLQQRKPKRSGREMERONLINACTCKMPKPPSPITPDLPLRASPSH	480						
QY	812	IPSVVVLPIITQGGYQHEHYVQPKKSEVAQALDEQAAITLEYKTIKEHLSKSPHNGVNLV	871						
Db	481	IPSVVVLPIITQGGYQHEHYVQPKKSEVAQALDEQAAITLEYKTIKEHLSKSPHNGVNLV	540						
QY	872	ENLDSLPPKYPORASLGGPPGASLSQTLGSKRLMHSSSYGVYKRSYTNLSITRSHQA	931						
Db	541	ENLDSLPPKYPORASLGGPPGASLSQTLGSKRLMHSSSYGVYKRSYTNLSITRSHQA	600						
QY	932	TYLKRNTSSNSGSHLSRNSFGGNDPPPPAPQVDSIQVHSSQPSQAVTVSHQPSLNA	991						
Db	601	TYLKRNTSSNSGSHLSRNSFGGNDPPPPAPQVDSIQVHSSQPSQAVTVSHQPSLNA	660						
QY	992	YNSLTRSGKLTPELKPDPKPSFAPLSTSMKENDACT	1030						
Db	661	YNSLTRSGKLTPELKPDPKPSFAPLSTSMKENDACT	699						

RESULT 7
AAB23043
ID AAB23043 standard; Protein, 630 AA.

XX AAB23043;
DT 16-JAN-2001 (first entry)

DE Human semaphorin protein-like splice variant, SECK pCR2.1-2864933.

XX SECK protein: human; secreted; membrane-associated; cancer;
XX Proliferation regulator; differentiation regulator; non-malignant tumour;
XX immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;
XX infection; inflammatory disorder; arthritis; haematopoietic disorder;
XX skin disorder; cardiovascular disorder; atherosclerosis; restenosis;
XX neurological disease; Alzheimer's disease; trauma; wounding;
XX spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
XX anti-HIV; anti-inflammation; antiatherosclerotic;
XX neuroprotective; vulnerability; antiallergic; antimicrobial; cardiac;
XX dermatological; gene therapy.

OS Homo sapiens.

XX	WO200053742-A2.
PN	
XX	14-SEP-2000.
PD	
XX	
XX	09-MAR-2000; 2000WO-US06280.
PF	
XX	
PR	09-MAR-1999; 99US-0123667.
PR	08-MAR-2000; 2000US-0123667.
XX	
PA	(CURAGEN CORP.
PI	
PI	Shinkets RA;
DR	WPI: 2000-594318/56.
DR	N-PSDB; AAA93630.
XX	
XX	Novel human membrane associated or secreted polypeptides and
PT	polynucleotides useful for diagnosis, prevention and treatment of
PT	pathological states such as cancer, immune, cardiovascular and
PT	neurological disorders
PS	Claim 1; Fig 15; 151pp; English.
XX	
XX	Sequences AAB23029-B23048 represent human SECK proteins. The SECK
CC	proteins of the invention are either secreted or membrane-associated
CC	proteins and act as regulator of cellular proliferation and
CC	differentiation. SECK proteins or nucleotides are useful for diagnosing
CC	the presence of, or predisposition to, a disease associated with altered
CC	levels of SECK proteins and nucleotides. The SECK proteins are also
CC	useful to screen compounds that modulate SECK activity or expression. The
CC	interaction of a SECK protein with other cellular proteins may be useful
CC	to modulate the activity of a partner protein. SECK nucleotides are useful
CC	cellular differentiation and cell survival. SECK nucleotides are useful
CC	for the recombinant expression of SECK protein, and may be used detect
CC	SECK mRNA or genetic lesions in the SECK gene. They may also be used to
CC	modulate SECK expression (e.g., using antisense oligonucleotides). SECK
CC	nucleic acid sequences are also useful for identifying a cell or tissue
CC	type in a biological sample, and in forensic biology. SECK primers or
CC	probes are useful for detecting the presence of SECK nucleotides and for
CC	screening tissue cultures for contamination. Diseases that may be treated
CC	or prevented using SECK proteins or nucleotides include cancer (e.g.,
CC	colorectal carcinoma, prostate cancer), benign tumours, immune disorders
CC	(including autoimmune diseases, transplant rejection, allergies, AIDS),
CC	infections, inflammatory disorders, arthritis, haematopoietic disorders,
CC	skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
CC	neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
CC	surgical or traumatic wounds, spinal cord injury), and skeletal
CC	disorders.
XX	
SO	Sequence 630 AA;
	Query Match 61.9%; Score 3373; DB 21; Length 630;
	Best Local Similarity 99.8%; Pred. No. 3.8e-284;
	Matches 628; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	17 GAGPPDESEPISSHGNTKYQYPVFGHKGRNTTQHRRLDIOQIMTMMGTLTYAARDHI 76
Db	1 GSGPDESEPISSHGNTKYQYPVFGHKGRNTTQHRRLDIOQIMTMMGTLTYAARDHI 60
QY	77 YTVDDITSHTEIYCSKKLTWKSQADVDYTRMGKHKDECHNFIYLLKKNDALFVCG 136
Db	61 YTVDDITSHTEIYCSKKLTWKSQADVDYTRMGKHKDECHNFIYLLKKNDALFVCG 120
QY	137 TNAFNPSGRNRYKMTLEPFGEFSGMARCPYDAKHANVALFADGKLYSATVTDPLAIDAV 196
Db	121 TNAFNPSGRNRYKMTLEPFGEFSGMARCPYDAKHANVALFADGKLYSATVTDPLAIDAV 180
QY	197 IYRSIGSEPTLRATYHDSKMLKEPFYQAVDYGDIYFFPELAEVNTMGKVPYPAQ 256
Db	181 IYRSIGSEPTLRATYHDSKMLKEPFYQAVDYGDIYFFPELAEVNTMGKVPYPAQ 240
QY	257 VCKNDMGSGQRVLEKQWTFPLKARLNCVSPGDSHRYFNILQAVTVIRNGRDVLAIFS 316

```

Db      241 VCNNDGSGRVLVEKQKTSFLKARLNGSVGDSHFYFNILQAVTDVLRINGRDVLAITS 300
Qy      317 TPVNSIFGSAVCAYMDLDSVFTGRKEQKSPDSTMTVPDDEKVPKRRGGCCAGSSSLE 376
Db      301 TPVNSIFGSAVCAYMDLDSVFTGRKEQKSPDSTMTVPDDEKVPKRRGGCCAGSSSLE 360
Qy      377 RYANSNEPDDTLNFITHTPLMDRAVPSIFNRPMFLRTMYRLLTKIADVTAAAGPYQNH 436
Db      361 RYANSNEPDDTLNFITHTPLMDRAVPSIFNRPMFLRTMYRLLTKIADVTAAAGPYQNH 420
Qy      437 VVFLGSEKGIILKFLARIGNSGFNLDSLPLEMSVYNSEKSYDGVEDKRIKMGQIDRAS 496
Db      421 VVFLGSEKGIILKFLARIGNSGFNLDSLPLEMSVYNSEKSYDGVEDKRIKMGQIDRAS 480
Qy      497 SSLVYASTCVIKVPLGRCERHCKCKTICASRDPIYGMKEGACSHLSPNSRLTEOD 556
Db      481 SSLVYASTCVIKVPLGRCERHCKCKTICASRDPIYGMKEGACSHLSPNSRLTEOD 540
Qy      557 IERNTDGLGDCNHSFVALNGHSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDST 616
Db      541 IERNTDGLGDCNHSFVALNGHSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDST 600
Qy      617 DPLGAVSSHNHODKKGVIRESYLKGHDQL 645
Db      601 DPLGAVSSHNHODKKGVIRESYLKGHDQL 629

RESULT 8
AAB95317
ID      AAB95317 standard; Protein; 574 AA.
XX      AAB95317;
DT      26-JUN-2001 (first entry)
XX      Human protein sequence SEQ ID NO:17568.
DE      Human
XX      Human; primer; detection; diagnosis; antisense therapy; gene therapy.
KW      Homo sapiens.
OS      Homo sapiens.
PN      EP1074617-A2.
XX      07-FEB-2001.
PD      28-JUL-2000; 2000EP-0116126.
XX      29-JUL-1999; 99QP-0248036.
PR      27-AUG-1999; 99QP-0300253.
PR      11-JUN-2000; 2000JP-0118776.
PR      02-MAY-2000; 2000JP-0183767.
PR      09-JUN-2000; 2000JP-0241899.
XX      (HELI-) HELIX RES INST.
PA      Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX      WPI; 2001-318749/34.
XX      Primer sets for synthesizing polynucleotides, particularly the 5602
PT      full-length cDNAs defined in the specification, and for the detection
PT      and/or diagnosis of the abnormality of the proteins encoded by the
PT      full-length cDNAs -
PS      Claim 8; SEQ ID 17568; 2537pp + CD ROM; English.
XX      The present invention describes primer sets for synthesizing 5602
CC      full-length cDNAs defined in the specification. Where a primer set
CC      comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC      to the complementary strand of a polynucleotide which comprises one of
CC      the 5602 nucleotide sequences defined in the specification, where the

```

```

CC      oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC      of an oligonucleotide comprising a sequence complementary to the
CC      complementary strand of a polynucleotide which comprises a 5'-end
CC      sequence and an oligonucleotide comprising a sequence complementary to a
CC      polynucleotide which comprises a 3'-end sequence, where the combination of
CC      oligonucleotide comprises at least 15 nucleotides and the combination in
CC      the 5'-end sequence/3'-end sequence is selected from those defined in
CC      the specification. The primer sets can be used in antisense therapy and
CC      in gene therapy. The primers are useful for synthesizing polynucleotides,
CC      particularly full-length cDNAs. The primers are also useful for the
CC      detection and/or diagnosis of the abnormality of the proteins encoded by
CC      the full-length cDNAs. The primers allow obtaining of the full-length
CC      cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC      AAH13633 to AAH18742 represent human amino acid sequences; AAH92446 to
CC      AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC      represent polynucleotides, all of which are used in the exemplification
CC      of the present invention.
XX      Sequence 574 AA;
SQ

```

```

Qy      398 MDEAVPSIFNRPMFLRTMYRLLTKIADVTAAAGPYQNHVFLGSEKGIILKFLARIGNS 457
Db      1 MDEAVPSIFNRPMFLRTMYR-----
Qy      458 GFINDSLPLEMSVYNSEKSYDGVEDKRIKMGQIDRASSLVYASTCVIKVPLGRCER 517
Db      21 -----CSYDGVEDKRIKMGQIDRASSLVYASTCVIKVPLGRCER 61
Qy      518 HGKCKTICIASRDPIYGMKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFVALNG 577
Db      62 HGKCKTICIASRDPIYGMKEGACSHLSPNSRLTFEODIERGNTDGLGDCNHSFVALNG 121
Qy      578 HSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDSTDPLGAVSSHNHODKKGVIRE 637
Db      122 HSSSLPSTTSDTAQEGYERSGMLDMKHLDSPDSTDPLGAVSSHNHODKKGVIRE 181
Qy      638 YLKGHDQLVPVTLIAIAVILAFVNGAVNSGITTYVCVCHRRKDVAVVQKREKELTHSRG 697
Db      182 YLKGHDQLVPVTLIAIAVILAFVNGAVNSGITTYVCVCHRRKDVAVVQKREKELTHSRG 241
Qy      698 SMSVYTKLSGLFGDTQSKDPPEALITPLMNGKLATPGNTAKMLIKADQHHLDLTALEPT 757
Db      242 SMSVYTKLSGLFGDTQSKDPPEALITPLMNGKLATPGNTAKMLIKADQHHLDLTALEPT 301
Qy      758 PESTPTLQOKRKPSGRSEMERONLINACTKMPMGSPVPTDLPPLASPSHIPSVVV 817
Db      302 PESTPTLQOKRKPSGRSEMERONLINACTKMPMGSPVPTDLPPLASPSHIPSVVV 361
Qy      818 LPITQGGYQHEVYVDQPKSEVAQMALEDOAAALEYKTKIHEHSSKSPNNGVNLVENDSL 877
Db      362 LPITQGGYQHEVYVDQPKSEVAQMALEDOAAALEYKTKIHEHSSKSPNNGVNLVENDSL 421
Qy      878 PKVPOREASLGPFGASLSQGTGSKRLEHMHSSSYGVYKRSYPTNLSLRSHQATTLKEN 937
Db      422 PKVPOREASLGPFGASLSQGTGSKRLEHMHSSSYGVYKRSYPTNLSLRSHQATTLKEN 481
Qy      938 NTSSNSHLSRNQSGFGDNPAPQVDSIQVHSSQSGQAVYVSRPSINAYNSLTP 997
Db      482 NTSSNSHLSRNQSGFGDNPAPQVDSIQVHSSQSGQAVYVSRPSINAYNSLTP 541
Qy      998 SGLKRTPSLKPDPVPPKSPFAPLSTSMKPDACT 1030
Db      542 SGLKRTPSLKPDPVPPKSPFAPLSTSMKPDACT 574

```

```

RESULT 9
AAB94104
ID      AAB94104 standard; Protein; 562 AA.
XX

```

AC AAB94104;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:14328.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 14328; 2537JP + CD ROW; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 562 AA;
 Query Match 54.2%; Score 2952; DB 22; Length 562;
 Best Local Similarity 100.0%; Pred. No. 1.4e-247;
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 SDSTAQGEYESSRGMLDWKHLSDPSDTPDLGAVSSHNDKKGVIRESLYKGDQLVV 648
 DB 121 SDSTAQGEYESSRGMLDWKHLSDPSDTPDLGAVSSHNDKKGVIRESLYKGDQLVV 180
 QY 649 TLLAIAVILAFVNGAVPSGITYVCVCDHRKRDVAIVQKKEKELTHSRGSMSSYTKLSGL 708
 DB 181 TLLAIAVILAFVNGAVPSGITYVCVCDHRKRDVAIVQKKEKELTHSRGSMSSYTKLSGL 240
 QY 709 FGDTSQCDPRPEALITPTLMNGKLAATPGAKMLIKADQHLDLTALPTPESTPTLQQR 768
 DB 241 FGDTSQCDPRPEALITPTLMNGKLAATPGAKMLIKADQHLDLTALPTPESTPTLQQR 300
 QY 769 KPSRGSREMERNOULINACTKMPMGSPVLPDLPASPSHIPSVVLPITQOQYHE 828
 DB 301 KPSRGSREMERNOULINACTKMPMGSPVLPDLPASPSHIPSVVLPITQOQYHE 360
 QY 829 YVDPKXSEVAQVMALEDQAATLEYKTIKEHLSSSPNHNQVNLVNLSPKPVQREASL 888
 DB 361 YVDPKXSEVAQVMALEDQAATLEYKTIKEHLSSSPNHNQVNLVNLSPKPVQREASL 420
 QY 889 GPPGASISQTLGSKRLMHHSSSYGVDPKRSYPTNSLTRSHQATTIKNNNTSSNSHLS 948
 DB 421 GPPGASISQTLGSKRLMHHSSSYGVDPKRSYPTNSLTRSHQATTIKNNNTSSNSHLS 480
 QY 949 RNOSFRGDNPPAPORVDSIQVHSSQPSGQAVTVSRQPSINAVNSLTRSGIKRTPELXP 1008
 DB 481 RNOSFRGDNPPAPORVDSIQVHSSQPSGQAVTVSRQPSINAVNSLTRSGIKRTPELXP 540
 QY 1009 DVPPKPSFAPLSTSMKPNDACT 1030
 DB 541 DVPPKPSFAPLSTSMKPNDACT 562
 RESULT 10
 ID AAM93444
 XX AAM93444 standard; Protein; 562 AA.
 AC AAM93444;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3088.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94365.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 3088; 1380bp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

CC molecules have been determined. Primers for synthesising the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special method. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 562 AA;

Query Match 54.1%; Score 2947; DB 22; Length 562;
 Best Local Similarity 99.8%; Pred. No. 3.8e-247;
 Matches 561; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 469 MSVNSKCSYDGVEDKRIKMGQDRASSSLVAFSTVIVIPGRGRHGCKKCTIAS 528
 DB 1 MSVNSKCSYDGVEDKRIKMGQDRASSSLVAFSTVIVIPGRGRHGCKKCTIAS 60
 QY 529 RDPYCGWIKGAGCSHLSPNRSLTFEODIERGNTDGLGDCNHFVALNGHSSLLPSTTT 588
 DB 61 RDPYCGWIKGAGCSHLSPNRSLTFEODIERGNTDGLGDCNHFVALNGHSSLLPSTTT 120
 QY 589 SDSTAQEGYSGRGMMDKHLIDPSDTPDGLAVSSNNHQQKGYTRRSYKGDQVLPV 648
 DB 121 SDSTAQEGYSGRGMMDKHLIDPSDTPDGLAVSSNNHQQKGYTRRSYKGDQVLPV 180
 QY 649 TLLAVALIAFMGAVFSGITVYCVCDHRKDVAVQKKEKELTSRSGSSSYTKLSGL 708
 DB 181 TLLAVALIAFMGAVFSGITVYCVCDHRKDVAVQKKEKELTSRSGSSSYTKLSGL 240
 QY 709 FGDTQSKPKEPAITLPMHNGKLTAPNTAMLIKADQHLDLTALPTPESTPTLQQR 768
 DB 241 FGDTQSKPKEPAITLPMHNGKLTAPNTAMLIKADQHLDLTALPTPESTPTLQQR 300
 QY 769 KPSRSGREMNQNTINACTKMPMGSPVITDPLPLASPSHISVVLPTQGGYQHE 828
 DB 301 KPSRSGREMNQNTINACTKMPMGSPVITDPLPLASPSHISVVLPTQGGYQHE 360
 QY 829 YVDQPKSEVQAQALDDQALTEYKTIKELSSKSPNHGVNLENLDSLPKVPQREASL 888
 DB 361 YVDQPKSEVQAQALDDQALTEYKTIKELSSKSPNHGVNLENLDSLPKVPQREASL 420
 QY 889 GPPGASISQGLSKLEMHSSSYGVYKRSYPTNSLTFSHQATTLKNNNTSSNSHLS 948
 DB 421 GPPGASISQGLSKLEMHSSSYGVYKRSYPTNSLTFSHQATTLKNNNTSSNSHLS 480
 QY 949 RNQSGRGNPPAPQDVDSIQVHSGQPGQAVTVSRQSLNAYNSLTSGIKRTPTSLKP 1008
 DB 481 RNQSGRGNPPAPQDVDSIQVHSGQPGQAVTVSRQSLNAYNSLTSGIKRTPTSLKP 540
 QY 1009 DVEPKSPAPLSTSMKPNDACT 1030
 DB 541 DVEPKSPAPLSTSMKPNDACT 562

RESULT 11
 ABG04066
 ID ABG04066 standard; Protein; 863 AA.
 XX
 AC ABG04066;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #4057.
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 FN WO200175067-A2.

XX 11-OCT-2001.
 XX
 PD 30-MAR-2001; 2001WO-US08631.
 XX
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYS-) HYSO INC.
 XX
 XX Diemanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS68253.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 PS Claim 20; SEQ ID No 34425; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 863 AA;

Query Match 53.0%; Score 2891; DB 22; Length 863;
 Best Local Similarity 80.9%; Pred. No. 5.9e-242;
 Matches 583; Conservative 28; Mismatches 92; Indels 18; Gaps 8;

QY 321 STPGSAVAYMDLIDASFVTFGRFKQKSPDSTWTPVDERVYKPPGCCAGSSSLERYAT 380
 DB 150 STPGSAVAYMDLIDASFVTFGRFKQKSPDSTWTPVDERVYKPPGCCAGSSSLERYAT 209
 QY 381 SNEPPDDTLNFKTHPLMDEAVPSIFNRPMLRTWRYRLTKIADVTAAGPQNTVFL 440
 DB 210 SNEPPDDTLNFKTHPLMDEAVPSIFNRPMLRTWRYRLTKIADVTAAGPQNTVFL 269
 QY 441 GSEKGIILKFLARIGNSGFLNLSFLEMSVYNSKCSYDGVEDKRI-MGMQDRASSSL 499
 DB 270 GSEKGIILKFLARIGNSGFLNLSFLEMSVYNSK-KWSTAKPKRVVITILNPGASFCI 328
 QY 500 YVAFSTCV---IKYPLGRG--ERHGKCKCTCIASRDYPC--GMIKEGAGCSHLSPSNR 550
 DB 329 TIRETYCKRRKHTWCPYCTIOXH-FCPCHLSGKETLCRVTGKMKVADRDSESLPYAA 387
 QY 551 LTFEQDI-ERGNTDGLGDCNHFVALNGHSSLSLESTTSSTAQEGYSGRGMMDKHL 609
 DB 388 MLLAQGMARCKEELGITLAIHKRAVGNRTK---TGPASRSPSSPCPLGLKWQTL 442
 QY 610 LQSPDSTPLGAVSSNNHQQKGYTRRSYKGDQVPTLLAVALIAFMGAVFSGIT 669
 DB 443 FPRRLRMFGGRRRSQLEAQRVIRESLYKGDQVPTLLAVALIAFMGAVFSGIT 502

Db 797 WMSCPSSRSATSMSTWTSPPKXAPRWRRMTRPPHMSIRPSRNIISAARVPTMGXTLMRTW 856
 QY 875 DSLPPK-----VPGREASLGPFGAS-----LSQTGSKRLEMHSSSYGVDRKSY 920
 Db 857 TACPPKHSGRPPWPVPPCLRPVXASGKCTTPLPLGLTIR----- 899
 QY 921 PTNSLTRSHQAT 932
 Db 900 GATPRTRSRKAT 911

RESULT 13
 AAB92688
 ID AAB92688 standard; Protein; 507 AA.
 AC AAB92688;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:11073.
 DE Human protein sequence SEQ ID NO:11073.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 OS Homo sapiens.
 XX EPI074617-A2.
 PN EPI074617-A2.
 PD 07-FEB-2001.
 PF 28-JUL-2000; 2000EP-0116126.
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-DAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS Claim 8; SEQ ID 11073; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a sequence and an oligonucleotide comprising a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and the specification. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialized methods. AAH03166 to AAH13628 and AAH13633 to AAH1742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX Sequence 507 AA;
 SQ Query Match 48.1%; Score 2621.5; DB 22; Length 507;
 Best Local Similarity 90.2%; Pred. No. 7e-219; Indels 55; Gaps 1;
 Matches 507; Conservative 0; Mismatches 0

QY 469 MSYTNSEKSGYDVEKRIWGMQLDRASSLSYAFSTCVIKVPLGRGCRKCKTCTIAS 528
 Db 1 MSYTNSEKSGYDVEKRIWGMQLDRASSLSYAFSTCVIKVPLGRGCRKCKTCTIAS 60
 QY 529 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNGHSSLLPSTTT 588
 Db 61 RDPYCGWIKKEGACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNGHSSLLPSTTT 108
 QY 589 SDSTAGEGYSRGMGLMDKHLDSBDSSTDPGLGANSNNHODKGVIRSEYKGGHOLVPV 648
 Db 109 -----GVIRSEYKGGHOLVPV 125
 QY 649 TLLAIVILAFVWGAVFSGITVYCVCDHRRKDVAVVQRKEKELTHSRGSMSSVTKLSGL 708
 Db 126 TLLAIVILAFVWGAVFSGITVYCVCDHRRKDVAVVQRKEKELTHSRGSMSSVTKLSGL 185
 QY 709 FGDTQSKDPKPEALITPLMNGKLTATPNTAKMLIKADQHLDTALPTPESTPTLOQR 768
 Db 186 FGDTQSKDPKPEALITPLMNGKLTATPNTAKMLIKADQHLDTALPTPESTPTLOQR 245
 QY 769 KPSRGRWRERNONLINACTKMPMGSPVFTDPLASPSHIPSVVVLEITQGYQHE 828
 Db 246 KPSRGRWRERNONLINACTKMPMGSPVFTDPLASPSHIPSVVVLEITQGYQHE 305
 QY 829 YVDQPKMEVAQMALEDQAATLEYKTIKELSSKSPNNGVNLVENLDSLPRKVPQRESL 888
 Db 306 YVDQPKMEVAQMALEDQAATLEYKTIKELSSKSPNNGVNLVENLDSLPRKVPQRESL 365
 QY 889 GPPGASLSQGLSKRLEMHSSSYGVDRKSYPTNSLTRSHQATTLKRNNTSSNSHLS 948
 Db 366 GPPGASLSQGLSKRLEMHSSSYGVDRKSYPTNSLTRSHQATTLKRNNTSSNSHLS 425
 QY 949 RNQSGRGRGNPPPAQRVDSIQVHSSQPSGCAVATSRQPSLNAVNSLTRSGLRTPSLKP 1008
 Db 426 RNQSGRGRGNPPPAQRVDSIQVHSSQPSGCAVATSRQPSLNAVNSLTRSGLRTPSLKP 485
 QY 1009 DVPKPSFAPLSTSMKENDACT 1030
 Db 486 DVPKPSFAPLSTSMKENDACT 507

RESULT 14
 AAG63213
 ID AAG63213 standard; Protein; 1086 AA.
 AC AAG63213;
 XX 01-OCT-2001 (first entry)
 DT 01-OCT-2001 (first entry)
 XX Amino acid sequence of a human semaphorin-like polypeptide.
 DE Expressed sequence tag; EST; semaphorin-like protein; neuronal growth;
 KW spinal cord damage; neurodegenerative disease; genetic neuronal defect;
 KW immunological disorder; lymphocyte dysfunction; viral infection; cancer.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH Peptide 1..16
 FT Peptide /note= "signal peptide"
 FT Domain 671..699
 FT /note= "transmembrane domain"
 XX W0200153466-A1.
 XX 26-JUL-2001.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 23, 2003, 17:09:07 ; Search time 19 Seconds
(without alignments)
2293.694 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450
Sequence: 1 MRSEALLLYFTLHPAGAGF.....PKPSPAPLSTSMKPNDACT 1030

Scoring table: BLOSUM62
Gapop 10.0 , Gapex 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents: AA.*

1: /cgm2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgm2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgm2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgm2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgm2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgm2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2013.5	36.9	888	US-09-077-940A-4	Sequence 4, Appli
2	1897	36.6	887	US-09-077-940A-2	Sequence 2, Appli
3	1519.5	27.9	930	US-09-254-594-6	Sequence 6, Appli
4	1428.5	26.2	929	US-09-121-713D-3	Sequence 3, Appli
5	969.5	17.8	730	US-08-121-713D-58	Sequence 58, Appli
6	969.5	17.8	730	US-08-835-268-58	Sequence 58, Appli
7	969.5	17.8	730	US-09-060-692-58	Sequence 58, Appli
8	969.5	17.8	730	US-08-833-391-58	Sequence 58, Appli
9	969.5	17.8	730	US-09-060-610-58	Sequence 58, Appli
10	969.5	17.8	730	PCT-US94-10151A-58	Sequence 58, Appli
11	880	16.1	650	US-08-121-713D-60	Sequence 60, Appli
12	880	16.1	650	US-08-835-268-60	Sequence 60, Appli
13	880	16.1	650	US-09-060-692-60	Sequence 60, Appli
14	880	16.1	650	US-08-833-391-60	Sequence 60, Appli
15	880	16.1	650	US-09-060-610-60	Sequence 60, Appli
16	880	16.1	650	PCT-US94-10151A-60	Sequence 60, Appli
17	852.5	15.6	655	US-08-556-422A-3	Sequence 3, Appli
18	852.5	15.6	655	US-08-121-713D-54	Sequence 54, Appli
19	852.5	15.6	655	US-08-835-268-54	Sequence 54, Appli
20	852.5	15.6	655	US-09-060-692-54	Sequence 54, Appli
21	852.5	15.6	655	US-08-833-391-54	Sequence 54, Appli
22	852.5	15.6	655	US-09-060-610-54	Sequence 54, Appli
23	852.5	15.6	655	PCT-US94-10151A-54	Sequence 54, Appli
24	845	15.5	712	US-08-121-713D-64	Sequence 64, Appli
25	845	15.5	712	US-08-835-268-64	Sequence 64, Appli
26	845	15.5	712	US-09-060-692-64	Sequence 64, Appli
27	845	15.5	712	US-08-833-391-64	Sequence 64, Appli

28	845	15.5	712	4	US-09-060-610-64	Sequence 64, Appli
29	845	15.5	712	5	PCT-US94-10151A-64	Sequence 64, Appli
30	779.5	14.3	724	1	US-08-121-713D-62	Sequence 62, Appli
31	779.5	14.3	724	2	US-08-835-268-62	Sequence 62, Appli
32	779.5	14.3	724	3	US-09-060-692-62	Sequence 62, Appli
33	779.5	14.3	724	4	US-08-833-391-62	Sequence 62, Appli
34	779.5	14.3	724	5	US-09-060-610-62	Sequence 62, Appli
35	779.5	14.3	724	6	PCT-US94-10151A-62	Sequence 62, Appli
36	748	13.7	775	4	US-09-308-179B-1	Sequence 2, Appli
37	736.5	13.5	862	4	US-08-556-422A-2	Sequence 2, Appli
38	654.5	12.0	607	4	US-08-556-422A-4	Sequence 2, Appli
39	586	10.8	477	1	US-08-136-922-2	Sequence 2, Appli
40	539	9.9	425	4	US-08-556-422A-7	Sequence 7, Appli
41	462	8.5	295	4	US-08-556-422A-6	Sequence 6, Appli
42	437	8.0	634	3	US-09-041-235-2	Sequence 2, Appli
43	437	8.0	634	4	US-09-771-467C-2	Sequence 2, Appli
44	437	8.0	666	3	US-09-240-410-2	Sequence 2, Appli
45	356.5	6.5	606	3	US-09-041-236-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-09-077-940A-4
; Sequence 4, Application US/09077940A
; Patent No. 6576441
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru et al.
; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
; FILE REFERENCE: 0020-4426P
; CURRENT APPLICATION NUMBER: US/09/077,940A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-940A-4
Query Match          36.9%; Score 2013.5; DB 4; Length 888;
Best Local Similarity 45.3%; Pred. No. 1.7e-182;
Matches 424; Conservative 129; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLLYFTLHPAGAGFEDSEPPISISHGNYTKQYVGVGHKGGNTTOR--HRLDIQWIM 62
DB 12 ALLLLLLLGGAGHGFEDDEPPLSVAPRDYLNHYVFGSGGRKLTTPAEGADDLNIQVVL 71
QY 63 INNGTLYIARDHYIYTDIDTSHTEETYSCKKLTWKSQADVDTCRMKGAKHDECHNFIK 122
DB 72 RVNRTLFGRDNLRYRELEFPPTSTELRYQRKLTWRSQSDINVCRMKGQEGECRNFVK 131
QY 123 VLLKKNDALFVCGSTNAFNSCRNYKMDLEPPFGDESGMARCPYDAKHANVALPADQKL 182
DB 132 VLLRDESTLVCSSNAFNPVCANVSIDTLPQVEDNISGMARCPYDPGHANVALFSDQML 191
QY 183 YSAVYTDPLADDAVYIRSLGSPFLIRYKHSKMYKEEYFQADVGYIYFEFEIYAVE 242
DB 192 FTATVTFPLADDAVYIRSLGSPFLIRYKHSKMYKEEYFQADVGYIYFEFEIYAVE 251
QY 243 YNTMGKVVFPVAVQCKKDMGSGORVLEKQWTSFLKARLNSVPGDSHFYFNIIQAVTV 302
DB 252 FNYLEKVVSKYARVCKNDVGSFVLEKQWTSFLKARLNSVPGDSHFYFNIIQAVTV 311
QY 303 IIRNGRDVLAFTSTPNYSIFGSAYCAVMDIASVFTGRFKEQKSPDSTWTPVYDEKVP 362
DB 312 VSLGRFPVLAVFSTPNSIFGSAYCAVMDIASVFTGRFKEQKSPDSTWTPVYDEKVP 371
QY 363 KPRPCCGAGSSLEKYATNSNEPDQTLNFICTHPLMDAVSIFNRPFLRTMVFYRLTK 422
DB 372 RPRPCCAPGM--QYNASSALPDILNFVFKTHPLMDAVFSLGHPWILRLTKMHQHLTR 429
```

```

QY 423 IAVDTAAGYQNHVTVFLGSEKGIILKEFLAR--IGNSGFLNDSLLEEMSVYNSEKCYD 480
DB 430 VAVDVAGAGPWGQTVTVFLGSEAGTVLKFVLRPNASTSGTSLVLEEFETRRPRCGRP 489
QY 481 GVED--KRTMGQDLORASSSLVYAFSTCVIKVPLRCERHCKCKTKCIASRDPYCGWIK 538
DB 490 GGGGTGQRLLSLELDASGGLLAAPRCVAVPVAAQCQYSGCMKNCIGSQDPYCGWAD 549
QY 539 GGACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSSLPSTTSDSTAQEGYE 598
DB 550 -GSCIFLSTGCTAAAPFQDVSGASTGSLGDC----- 578
QY 599 SSGKMLDWKHLNLDSPDSTDPGLAVSSHNDQKGVIRBSYLGKHDQLYPVYTLAIAVILA 658
DB 579 -----TGILRASLSDDRAGLVSVNLVLTSSVAA 606
QY 659 FVMGAVFSGITV-YVCCHRRKRDVAVVOKEKE--LTHSRGSMSSVTLSGLFGDTQK 715
DB 607 FVVGAVVSGFVGMFGLRERRELA--RRKDEAILAHGAGEAVLSVSL---GERRAQ 660
QY 716 DP-----KPEAILTPLMNGKLAATPGNTAK-MLIKADQHLDLTALPTPEST 761
DB 661 GGGGCGGGGGGAGVAPPEALIAPIQNG-----WAKATLLQGGPHDLDSGLPTPEQT 713
QY 762 PTLQCKRP-----SRGSEMERONLIINACTKMPMGSPVPTDLP---LRASPEH 811
DB 714 P-LPQKRLPTPHPHALGPRAWDH-----GHPLPFASSSLILLIAPAR 757
QY 812 ISVVVLPITQGYGHEVVDQPKMSEVAQMALEDQAATLEYKTIKEHLSKSPNHQVNLV 871
DB 758 ABQPPAPPEPTPDGGLYIARPERASHGPFPLTPHASPERRKV-----SAPTGLDPA 811
QY 872 ENLDSLP-PKVPQREASL-----GPPGASLSQT 898
DB 812 SAADGLPRPMSPPPTGSLRRPLGPHAPPAATLRT 846

RESULT 2
US-09-077-940A-2
; Sequence 2, Application US/09077940A
; Patent No. 6576441
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru et al.
; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
; FILE REFERENCE: 0020-4426P
; CURRENT APPLICATION NUMBER: US/09/077,940A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2:
; LENGTH: 887
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-077-940A-2

Query Match 36.6%; Score 1997; DB 4; Length 887;
Best Local Similarity 42.7%; Pred. No. 6,3e-181;
Matches 433; Conservative 132; Mismatches 260; Indels 190; Gaps 25;

```

```

DB 190 GMLFATVTDPLAIDAVIYRSLGDRPILRTIVKHDSKMFKEYFYHAHEWGSHVYEFFREI 249
QY 240 AVEYNTMGKVVFPVAVQYCKNDMGQSOYLEKQMTSEFKARLNCVSPEDSHFYENILQAV 299
DB 250 AMEFNYLEKVVAVSRVAKCKNDVGSPVLEKQMTSEFLKARLNCVSPEDSHFYENILQAV 309
QY 300 TDVIRINRDVVLATFSPNYSIGSANCAVMDMLDIAVTFGRFPECKSPSTMTVPVDE 359
DB 310 TGVVSLGRPVLIAVFTSPNSISPGSAVCADIMOVAAVFGSRFPEQSPSITWTFVED 369
QY 360 RVKPRPGCCAGSSSLEBYATSNNEPDDTLNFIKTHPLMDAVPSIFNRPMFLRTMVRYR 419
DB 370 QVPRRPPGCCAABGM-QYNASNALPDEILNFVKTHPLMDAVPSIGNSPMIVRLIRHQ 427
QY 420 LTKIADVDTAAGYQNHVTVFLGSEKGIILKEFLAR--IGNSGFLNDSLLEEMSVYNSEK 477
DB 428 LTRVAVDVAGAGPWGQTVTVFLGSEAGTVLKFVLRPNASTSGTSLVLEEFETRRPRCG 487
QY 478 ---SYDVEDKRIKMGQDLORASSSLVYAFSTCVIKVPLRCERHCKCKTKCIASRDPYCG 534
DB 488 GRSSAGAGWGORLLSLELDASGGLLAAPRCVAVPVAAQCQYSGCMKNCIGSQDPYCG 547
QY 535 WKEGACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSSLPSTTSDSTAQ 594
DB 548 WAPD-GSCIFLSTGCTAAAPFQDVSGASTGSLGDC----- 580
QY 595 EGYESRGKMLDWKHLNLDSPDSTDPGLAVSSHNDQKGVIRBSYLGKHDQLYPVYTLAIA 654
DB 581 -----TGILRASLSDDRAGLVSVNLVLTSS 604
QY 655 VILAVMGAVFSGITV-YVCCHRRKRDVAVVOKEKE--LTHSRGSMSSVTLSGLFGDTQK 705
DB 605 SYAAVVGAVVSGFVGMFGLRERRELA--RRKDEAILAHGAGEAVLSVSLGRRGT 662
QY 706 --SGLFQDTQSKDPKPEAILTPLMNGKLAATPGNTAKMLIKADQHLDLTALPTPESTPT 763
DB 663 GTGGRGAGGCGGAPPEALIAPIQNG-----TKAALLHGGPHDLDSGLPTPEQTP 715
QY 764 LQCKRPKPR-----GSEMERONLIINACTK-----DMPMGSPVPTDLP--- 802
DB 716 LPQKRLPTTHPHALGPRAWDHSHALLSASASTLLIATRAPQPP-----VTESG 770
QY 803 -----LPLRASPSPHPSVVVLPITQGYGHEVVDQPKMSEVAQMALEDQAATLEYKTIK 856
DB 771 PPSRLCAPRSCASHGDEPLTP-----HASPDRRVVSAPTGLDSSVG----- 816
QY 857 EHLSSKSPNHGVNLVENDSLP-PKVPQREASL-----GPPGASLSQTGSKRLEMHS 909
DB 817 -----DDLPGPWSPPATSSLRPPGPHGPPYALALRT-----HT 849
QY 910 SSGYVDYKSYPTNLSLTRHQATTLKRNNTSNSSGHSLSRMSFGGD-NPPAP 963
DB 850 FMSG-----EAPGGRPRRHA-----PADSTHL--LPCGTGEXTAPVP 887

RESULT 3
US-09-254-594-6
; Sequence 6, Application US/09254594
; Patent No. 6566094
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
; FILE REFERENCE: 0020-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6:
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

NAME/KEY: misc feature
 LOCATION: ()
 OTHER INFORMATION: Tissue Type: Child Brain
 NAME/KEY: misc feature
 LOCATION: ()
 OTHER INFORMATION: Identification Method: P for resulting peptide
 us-09-254-594-6

Query Match 27.9%; Score 1519.5; DB 4; Length 930;
 Best Local Similarity 35.9%; Pred. No. 2,4e-135;
 Matches 381; Conservative 135; Mismatches 356; Indels 189; Gaps 32;

6 LLLYFTLLHPAGAFPEDESEPISSHGNTKQYVFNVGHKRGNTTORHRLDIOIMIMN 65
 13 LLLLSLPH-TQAAPDPLPLISDQTSPLSWFGLBEDDAVAEL-GIDFQRFLLTN 70
 66 GTIYIARDHIYVDIDTSHT-ETIYCSKLTWKSQADVDTCMKGKHDECHNFIVL 124
 71 RTLLVAARDHVSFDLOAEERGEGLVFNKFLTWRSQ--DMENCAVRGLTDECINYIRVL 128
 125 LKKNDALFYCGTNAFNPSCRNKMDTLEPFGEFSGMACPYDAKANAVALPADGKLYS 184
 129 VPNDSCILLACGINSFSPVCRSYGITSLOQGEBELSGQACPFATQSNVAIFABGSLYS 188
 185 ATVTDLAIDAVIYRSLGESPTLFTVGHDSKMLKEPYFQAVDGYITFFPEIAYEVN 244
 189 ATAADFQASDAVYVRSIGPQPLRSAXYDSKWLREPHFVALBHGHDVFEFFREVSEDA 248
 245 TMGVVPPRYAOYCKNDMGSGQVLEKQWTSFLKARLNCSPVGDSHFYFNILQAVTVIR 304
 249 RLGRVQFSRVARVCKRMGSPRALDRHTSFLKRLNCSVPDSTYFVLOALTGPNV 308
 305 INGRDVLAATFSTYNSIPGSAVCAVMDLADIAVFTGRFKEQKSPDSTWTPVDERVCKP 364
 309 LHGSALFGVTTQTGNSIPGSAVCAVFLDEIERGFEKQKQSLDGMATVSDRVPSP 368
 365 RPGCCAGSSSLERYATSNPEPDDTLNFKTHPLMDAIVPSIFNRPWFLRTMVRRLIKIA 424
 369 RFGSCAGVGAALFSSSRDLDDVLTFLKAPLLDPAPVPTHOP-LTLTTSRALLTQVA 427
 425 VDTAAGPQNTTVFLGSEKGIILKFLARIGNSGFLNLSLEMSVYNSSEKSCAYDVED 484
 428 VDNAGAPHSNTVFLGSDGTFLKVLTPGSRSG-PEPILLEIDAYSPPRCGSKRTAQ 486
 485 --KRIIMQDLRASSSLVAFSTCVIKVPLGRCEHCKKXTCIASRDPYCGMIKEGAC 542
 487 TARRIIGELDTGHRFLFVAFSGCIVYPLRSCARHAGCQSCIASQPYCGMSSRS-C 545
 543 SHLSPNSHLFEQDIERGNTDGL-GDCHNSFVALNGHSSSLPSTTSDTAQEGYSR 600
 546 VDIRSGGTVDQ--AGNOSMEHGDQDG-----ATGSOQSGRGSAY--- 586
 601 GGMIDMKLLDSPOSTDPLGAVSSHNHODKGVIRESLYKGDHCLVPTLLAIATLAFV 660
 587 -----GVRDLPPLASASRSVPLILLASIAAFA 615
 661 MGAVFSGITVYCVCDHRRKDAVAVQREKELTHSRG-----SMSSVTKLSGL 708
 616 LGASVSGILLVSAC--RR-----AHRRRKDIETPGPRLPLSLARLHG- 659
 709 FGDTSQDKPP--EAILPLMHNGKLAFTGNTAKMLIKADQHLHLLTALPTPESTPTLQ 766
 660 -GGPPPPPSKDGDAVQTPOLYTTFLPPEGVPP-----ELACLPPESTPELAV 709
 767 KKPGRSGSEMERONLINAACKMDPMPGSPVIFTDLPLRASPSH-----IPSVVLLPT 821
 710 KHLRAAGD-PWEMNQRRNA-----KEGPRSGGHAAGPARAVLVRP-- 752
 822 QCGYQHEYVDQPMSEVAQMALEDQAATLEYKTIKEL-----SSKSPNHGVLVENLDS- 876
 753 -----PPGCGPG-----AVEVTLLELLRYLHGQPPRPGAGAPPAPLTSR 793
 877 -LPKVPQREASLGPFGASISOTGLSKLEMHSSSYGVYKRSYVTNLSLTHSQATLX 935

794 ALPPE--PAPALLGGSPRPHCASPRLRDV-----PPRGACASAPA----- 833
 936 RNNINSSNSHL-----SRNOSFRGDNPPAPQRVDISIQHSSQPGCAVTVSRQS--- 988
 834 -PPALGAPAPRLGVGGGRLLPFGHAPAPALLTRV-----PFGGSPRYSGGPGKHL 883
 989 --LMAINSILTRGILKRTPSLKPDPVPEKPSF-APLSTSMKEN 1026
 884 LVIGRPEGYRGALKKRVDEKPOLSLKPLVPBSSQAVPN 924

RESULT 4

US-09-254-594-3
 ; Sequence 3; Application US/09254594
 ; Patent No. 6566094
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMURA, Toru
 ; APPLICANT: KIKUCHI, Kaoru
 ; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
 ; FILE REFERENCE: 0020-4527P
 ; CURRENT APPLICATION NUMBER: US/09/254,594
 ; CURRENT FILING DATE: 1999-05-11
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 929
 ; TYPE: PRI
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: ()
 ; OTHER INFORMATION: Tissue Type: Brain
 ; NAME/KEY: misc feature
 ; LOCATION: ()
 ; OTHER INFORMATION: Identification Method: P for resulting peptide
 US-09-254-594-3

Query Match 26.2%; Score 1428.5; DB 4; Length 929;
 Best Local Similarity 34.0%; Pred. No. 1.1e-126;
 Matches 360; Conservative 147; Mismatches 355; Indels 197; Gaps 32;

6 LLLYFTLLHPAGAFPEDESEPISSHGNTKQYVFNVGHKRGNTTORHRLDIOIMIMN 65
 11 LLLLSLPHQAQTAFPDPLPLISDQTSPLSWFGLBEDDAVAEL-GIDFQRFLLTN 69
 66 GTIYIARDHIYVDIDTSHT-ETIYCSKLTWKSQADVDTCMKGKHDECHNFIVL 124
 70 RTLLVAARDHVSFDLOAEERGEGLVFNKFLTWRSQ--DMENCAVRGLTDECINYIRVL 127
 125 LKKNDALFYCGTNAFNPSCRNKMDTLEPFGEFSGMACPYDAKANAVALPADGKLYS 184
 128 VPNDSCILLACGINSFSPVCRSYGITSLOQGEBELSGQACPFATQSNVAIFABGSLYS 187
 185 ATVTDLAIDAVIYRSLGESPTLFTVGHDSKMLKEPYFQAVDGYITFFPEIAYEVN 244
 188 ATAADFQASDAVYVRSIGPQPLRSAXYDSKWLREPHFVALBHGHDVFEFFREVSEDA 248
 245 TMGVVPPRYAOYCKNDMGSGQVLEKQWTSFLKARLNCSPVGDSHFYFNILQAVTVIR 304
 249 RLGRVQFSRVARVCKRMGSPRALDRHTSFLKRLNCSVPDSTYFVLOALTGPNV 308
 305 INGRDVLAATFSTYNSIPGSAVCAVMDLADIAVFTGRFKEQKSPDSTWTPVDERVCKP 364
 309 LHGSALFGVTTQTGNSIPGSAVCAVFLDEIERGFEKQKQSLDGMATVSDRVPSP 368
 365 RPGCCAGSSSLERYATSNPEPDDTLNFKTHPLMDAIVPSIFNRPWFLRTMVRRLIKIA 424
 369 RFGSCAGVGAALFSSSRDLDDVLTFLKAPLLDPAPVPTHOP-LTLTTSRALLTQVA 427
 425 VDTAAGPQNTTVFLGSEKGIILKFLARIGNSGFLNLSLEMSVYNSSEKSCAYDVED 484
 428 VDNAGAPHSNTVFLGSDGTFLKVLTPGSRSG-PEPILLEIDAYSPPRCGSKRTAQ 486
 485 --KRIIMQDLRASSSLVAFSTCVIKVPLGRCEHCKKXTCIASRDPYCGMIKEGAC 542
 487 TARRIIGELDTGHRFLFVAFSGCIVYPLRSCARHAGCQSCIASQPYCGMSSRS-C 545
 543 SHLSPNSHLFEQDIERGNTDGL-GDCHNSFVALNGHSSSLPSTTSDTAQEGYSR 600
 546 VDIRSGGTVDQ--AGNOSMEHGDQDG-----ATGSOQSGRGSAY--- 586
 601 GGMIDMKLLDSPOSTDPLGAVSSHNHODKGVIRESLYKGDHCLVPTLLAIATLAFV 660
 587 -----GVRDLPPLASASRSVPLILLASIAAFA 615
 661 MGAVFSGITVYCVCDHRRKDAVAVQREKELTHSRG-----SMSSVTKLSGL 708
 616 LGASVSGILLVSAC--RR-----AHRRRKDIETPGPRLPLSLARLHG- 659
 709 FGDTSQDKPP--EAILPLMHNGKLAFTGNTAKMLIKADQHLHLLTALPTPESTPTLQ 766
 660 -GGPPPPPSKDGDAVQTPOLYTTFLPPEGVPP-----ELACLPPESTPELAV 709
 767 KKPGRSGSEMERONLINAACKMDPMPGSPVIFTDLPLRASPSH-----IPSVVLLPT 821
 710 KHLRAAGD-PWEMNQRRNA-----KEGPRSGGHAAGPARAVLVRP-- 752
 822 QCGYQHEYVDQPMSEVAQMALEDQAATLEYKTIKEL-----SSKSPNHGVLVENLDS- 876
 753 -----PPGCGPG-----AVEVTLLELLRYLHGQPPRPGAGAPPAPLTSR 793
 877 -LPKVPQREASLGPFGASISOTGLSKLEMHSSSYGVYKRSYVTNLSLTHSQATLX 935

```

REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TEXT:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-121-713D-58

Query Match 17.8%; Score 969.5; DB 1; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29

      8 LYPTLLHPAGAGFPEDESPISISHGNVTKQYVPVGHKKRGENTTOR-----HRLDIQMI 61
      11 LLWVAIH-AAAVNDVSP-----KMYVQF-----GERRVGRFLGNESHKHDFKLL 54
      62 MINAGTLYIARDDHLYTWDID--TSHTTEIYCSKLLTWKSPQADVTORMKGKHDECHN 119
      55 EKDNHSLVIGARNIYVYNISLRDLTEFTQ-----RIEHHSSGAHRELTYLKGSSEDDCN 109
      120 FIKVLLKNDNALFVCGTNAEPSCRANKMDLBEFGD-----EPGMAKCEPYDAKHAHV 174
      110 YIRVLAKIDDDRVLLICGTNAKPLCRHNAIKD-----GDYVKEKEYGRLCFDPDHNST 165
      175 ALFADGKLYSATVTDFLAIDAVIYRSLGESPTLRTVKHDSKYLKEPYFQAVDYGVIYF 234
      166 AIVSGQGLYSATVADFGSGDPLIRG-----PLRTERSDLKQIANPNTMEYNDPIEF 220
      235 FPRELAVENYMGKVIYFPRVAVQCNKDMDGSGQRLVEKQMTSLTKARLNCSEVGDSHFYFN 294
      221 FPRELAVENYMGKVIYRVARVAVCHDKGPHQGGDR-WTSLKSLKLNCSVPDDYEFYFN 279
      295 ILQAVTDVIRLN-GRDV--VLATFSTPYNSIPGSANVAYMDLIDIASVFTGRFKQKSPD 350
      280 EIQSTSDIIEGNYGQVEKLIYGVFTTPYNSIIGSAVCAFSKMSLIESEFGPEKQOETMN 339
      351 STWTPVDPDERVYKPRPGCCAGSSSLERVATSNBPDDTLNFIKTHPLMEAVYSINRPW 410
      340 SMLVAVPSLKVPEPRPGQVND-----SKTLDPVSNVYKSHMLDAAVAFPTRTI 390
      411 FLRTWVRKYLTKIAND-----ZAAQPYQNTYVFLGSEKGIILKFLARIGNSGFLN----- 461
      392 LIRISIQYVFTKIANDQQVTRPDG--KADVLFIGIDGQVIAKAL-----NSASFDSQDV 445
      462 DSLFLEEMVYVNSEKSYDGVEDKRIWQMOLDRASSSLVAVASTVIKYPVLRG--ERHG 519
      446 DSVVIBELQVLP-----PGVYKRLYYVRMDDDSKLVVSDDELTIAIKRIGSGDKIT 499
      520 KCKKTCIASRDPYCGMIKGGACSH--SPN---SKLTFFQDIERGNTDGLGDCNPSVAL 575
      500 NCRF-CVSLQDPYCAMDNVELKCTAVGSDPWSAGKRRFQNISLGHKACGAPROTEIV- 557
      576 NGHSSLLISTTSDSTA-----OEGVESRCGLMDWKILIDSPSTDPLGAVSSHNQ 628
      558 ---ASPVLPQPTTKSGDPVSHIQAEPFPE--IDNEIVIGVDSSNVIPNTLAEINHA 611
      629 DKGVIRESLYKGDHDLVPV---TL-LAI-----AVIIAPYMGAVFS 666
      611 GSK-----LPSSQEKPLPYTAETLTIAIVTSCLGALVVGFTSGFLFS 652

RESULT 6
US-08-835-268-58
; Sequence 58, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.

```



```

APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,268
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/06/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-835-268-58

Query Match      17.8%; Score 969.5; DB 1; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LFTLLFAGAGFPEDESPISISHGNYTKQYVFVGHKPGENTQR-----HRLDIQMI 61
DB 11 LLMVALH-AAAMNDVSP-----KMYVGF-----GERVQRFIGNESHKHQFKLL 54
QY 62 MINGGTLYIARHIIYTVDD--TSHTEIYCSKLTWKSKQADVDTCRMKSKHDEKH 119
DB 55 EKDHNSILVGARNTVYVNI SLRDLTEFTQ-----RIEWHSSGAHRELCTYLGKSEDDCN 109
QY 120 FIKYLLKNDLALFVCGFNAPNSGRNYKMDTLEPGD-----EFSGMARCPYAKHANY 174
DB 110 YIRVIAKIDDDRVLICGTNAIKPLCRHIALD-----GIYVVEKEVEGGLCPFPBDNST 165
QY 175 ALFADGKLYSATVTDPLAIDAVIYKSLSEPTLRTVKHDSKMLKEPFVQAVDYGYIYF 234
DB 166 AIYSGQIYSATVADFSGTDPLIYRG-----PLRTERSDLKQINAPNFVNTMEYVDFEYF 220
QY 235 FFEREIAVEYNTMGKVFPRVAVQVCKNDMGSGORVLEKQWTSFLKRLKLCVPGDSHEFYN 294
DB 221 FFEREIAVEYINCGKAIYGRVAVVCHDKGPGHGGDR-WTSFLKRLKLCVPGDPPEYFN 279
QY 295 ILQAVTVIRLN-GADV---VLATSTPYNSIPGSAYCAVDMLDIASVFTGRFKEQKSPD 350
DB 280 EIQSTSDIIEGNYGQVEKLIYGVETTPNLSIGSAYCAFSKSLLESPDGFKEQETMN 339
QY 351 STWTVDPDERVYKPPGCCAGSSSLERYATSNFPDDTLNFKTLPMLDEAVESIFPNRW 410
DB 340 SNMVLAVPSLKVPERPGQCVD-----SRLPDVSYNFEVKSHTLMDAFAFAFTPI 391
QY 411 FLRTWVRIRLTKIAND---TAAGPYQNHITVFLQSEKIIILKFLARIGNSGFLN----- 461

```

```

DB 392 LRISIQRYFTIANDQVTRPDG--KAYDVLFIGDDDKVIAL-----NSASFDSDDTV 445
QY 462 DELFLEMSVYNSEKSYDGEDKRIQMOLDRASSLVYASTCYIKVPLGIC--ERHG 519
DB 446 DSVIIEELQVLP-----PGVPVKLYVRMGDDSKLVVSDDELIAIKLHRCGSDKIT 499
QY 520 KCKKTTIARSDRYCGMIKRGACNSH--SPN-----SRLTEQDIERGNTDGLGCHNSFVAL 575
DB 500 NCRE-CVSLQDPYCAWDNVELKCTAVGSPDWSAGKRRTFQNTSLGHHKACGRRPQTEIV- 557
QY 576 NGHSSSLPSTTSDSTA-----QEGYESRGMDWMHLLDSDPSTPLGAVSSHNQ 628
DB 558 ---ASPVTQPTTKSGDPVSHINQAEFPE---IDNEIVIGVDDSNVLPNTLAELINHA 610
QY 629 DKKGVIRESYLKGHDLVPPV---TL-LAI-----AVIIAPMGVFS 666
DB 611 GSK-----LPSSGKELPIYTAETLTIAIVISCLALVGVFISGLIFS 652

```

```

RESULT 7
US-09-060-692-58
Sequence 58, Application US/09060692
Patent No. 5935865
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,692
FILING DATE:
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-692-58

Query Match      17.8%; Score 969.5; DB 2; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LFTLLFAGAGFPEDESPISISHGNYTKQYVFVGHKPGENTQR-----HRLDIQMI 61

```

```

Db      11 LLMVALH-AAAMVNDVSP-----KMYVOF-----GERRVORFLGNESHKHDFKLL 54
QY      62 MIMNGTLYIARDDIYTVDID--TSHTTEIYCSKKLTWKSROADVDTRMKGKHEDCHN 119
Db      55 EKDHNSLIVGARNIYINISLRLTFTEQ-----RIEHSAGAHRELCTYLGKSDDCQN 109
QY      120 FIKVLLKKNDDALFVCGTNAFNSCRNYKMDTLEPFGD-----EFSGARCPYDAKHANV 174
Db      110 YIRVLAKIDDDRVLLCGTNAKPELCRHYALKD-----GDYVEKEVEYEGGLCPFDPHNST 165
QY      175 ALFADGKLYSATVTPFLAIDAVIYSLGESPTLRTVKHDSKMLKEPYVOAVDGYDYTYF 234
Db      166 AYSSEGOLYSATVADPFGSDPLIYRG-----PLKTERSDLKQLNAPNVTNMEYNDFTLF 220
QY      235 FPREIAVENYTMGKVVPRPAQVCKNDMGSGQVLEKQMTSFLKARLNCVSGDSHFYFN 294
Db      221 FPREIAVEYINCGKAIYSKVARVACHDKGPHQGDGR--WTSFLKRLNCSVGDYFPYFN 279
QY      295 ILQAVTVIRIN-GRDV--VLATSTPYNSIPGSAVCAVMDLDAVFTGRFKEQKSPD 350
Db      280 EIQSTSDIIEGNYGGVEKLIYGVFTTPVNSIGSAVCAFSKMSILSEFDGPFKEQETWN 339
QY      351 STWTPVDEDRVPRKPRGCCAGSSSLERVATSNFPDDTINFIKTHPLMDEAVPSIFNRPW 410
Db      340 SNMLAVPSLKVEPRPGQCVND-----SRTLPPVSVNPFVKSHTLMDAVAPAFTRPI 391
QY      411 FLRTWRYRLTKIADV-----TAAGPYQNTVFLGSEKGIILKFLARIGNSGFLN----- 461
Db      392 LRISIQYFTKIAVQVQVTRPDG--KAYDVLFIGTDDGKVIKAL-----NSASFSSDTV 445
QY      462 DSLFLEEMSVYNSEKSYGVDEKRIWQOLDRASSLYVANSTCYIKVPLGRG--ERRG 519
Db      446 DSVVIELQVLP-----PGVPYKMLYVVRMDGDGSKLYVSDDELLAIKLRGSGDKIT 499
QY      520 KCKKTCIASRDPYCGWIKEGGACSHL--SPN-----SRLTFEODIERGNTDGLGDCHNSFVAL 575
Db      500 NCRE--CVSLQDPYCAMDNVELKCTAVGSPWSAGKRFFIONISLGHKACGGRPQTEIV- 557
QY      576 NGHSSSLPSTTSDSTA-----OEGVESRGMGLDWKHLDPSTPDLGAVSSHQ 628
Db      558 ----ASFPVQPTTKSGDPVSHIQAEFEPE--IDNEIVIGVDSNVIPNTLAIENHA 610
QY      629 DKKGVIRESYLKGDQLVPV---TL-LAI-----AVLIAFYMGAVFS 666
Db      611 GSK-----LPSSQEKLPITYTAETLITAIIVTSCLGALVVGFSIGFLFS 652

RESULT 8
US-08-833-391-58
; Sequence 58, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matches, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391

```

```

; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-391-58

Query Match      17.8%; Score 969.5; DB 3; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY      8 LYFTLHPAGAGFPEDSEPISSHGNTYKQVPEVGHKPGANTQR-----HRLDQMI 61
Db      11 LLMVALH-AAAMVNDVSP-----KMYVOF-----GERRVORFLGNESHKHDFKLL 54
QY      62 MIMNGTLYIARDDIYTVDID--TSHTTEIYCSKKLTWKSROADVDTRMKGKHEDCHN 119
Db      55 EKDHNSLIVGARNIYINISLRLTFTEQ-----RIEHSAGAHRELCTYLGKSDDCQN 109
QY      120 FIKVLLKKNDDALFVCGTNAFNSCRNYKMDTLEPFGD-----EFSGARCPYDAKHANV 174
Db      110 YIRVLAKIDDDRVLLCGTNAKPELCRHYALKD-----GDYVEKEVEYEGGLCPFDPHNST 165
QY      175 ALFADGKLYSATVTPFLAIDAVIYSLGESPTLRTVKHDSKMLKEPYVOAVDGYDYTYF 234
Db      166 AYSSEGOLYSATVADPFGSDPLIYRG-----PLKTERSDLKQLNAPNVTNMEYNDFTLF 220
QY      235 FPREIAVENYTMGKVVPRPAQVCKNDMGSGQVLEKQMTSFLKARLNCVSGDSHFYFN 294
Db      221 FPREIAVEYINCGKAIYSKVARVACHDKGPHQGDGR--WTSFLKRLNCSVGDYFPYFN 279
QY      295 ILQAVTVIRIN-GRDV--VLATSTPYNSIPGSAVCAVMDLDAVFTGRFKEQKSPD 350
Db      280 EIQSTSDIIEGNYGGVEKLIYGVFTTPVNSIGSAVCAFSKMSILSEFDGPFKEQETWN 339
QY      351 STWTPVDEDRVPRKPRGCCAGSSSLERVATSNFPDDTINFIKTHPLMDEAVPSIFNRPW 410
Db      340 SNMLAVPSLKVEPRPGQCVND-----SRTLPPVSVNPFVKSHTLMDAVAPAFTRPI 391
QY      411 FLRTWRYRLTKIADV-----TAAGPYQNTVFLGSEKGIILKFLARIGNSGFLN----- 461
Db      392 LRISIQYFTKIAVQVQVTRPDG--KAYDVLFIGTDDGKVIKAL-----NSASFSSDTV 445
QY      462 DSLFLEEMSVYNSEKSYGVDEKRIWQOLDRASSLYVANSTCYIKVPLGRG--ERRG 519
Db      446 DSVVIELQVLP-----PGVPYKMLYVVRMDGDGSKLYVSDDELLAIKLRGSGDKIT 499
QY      520 KCKKTCIASRDPYCGWIKEGGACSHL--SPN-----SRLTFEODIERGNTDGLGDCHNSFVAL 575
Db      500 NCRE--CVSLQDPYCAMDNVELKCTAVGSPWSAGKRFFIONISLGHKACGGRPQTEIV- 557
QY      576 NGHSSSLPSTTSDSTA-----OEGVESRGMGLDWKHLDPSTPDLGAVSSHQ 628
Db      558 ----ASFPVQPTTKSGDPVSHIQAEFEPE--IDNEIVIGVDSNVIPNTLAIENHA 610
QY      629 DKKGVIRESYLKGDQLVPV---TL-LAI-----AVLIAFYMGAVFS 666
Db      611 GSK-----LPSSQEKLPITYTAETLITAIIVTSCLGALVVGFSIGFLFS 652

```

```

RESULT 9
US-09-060-610-58
Sequence 58, Application US/09060610
Patent No. 6344544
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David R.
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESSES:
ADDRESS: SCIENCE & TECHNOLOGY IAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,610
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,268
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-610-58

Query Match      17.8% Score 969.5; DB 4; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LYFTLLHRAAGFPPDSSEPTISHSNNTKQYVFGHKGKNTTOR-----HRLDQMI 61
DB 11 LHMVALH-AAAWNDVSP-----KMYVQF-----GEERVQRELGESHKDFEL 54
QY 62 MIMMGTLYIARBDHYTDID--TSHTEIYCSKLTWKSROADVDTCKMGKHKDECHN 119
DB 55 EKDNHSLVAGARNIYINISLRDLTFTEO-----RIEMHSSAHNELCYLKGKSDDCN 109
QY 120 FIKVLLKKNDDALFVCGTNAFNPSCRNKMDTLEPFGD-----EFGMARCPYDAKMANV 174
DB 110 YIRVLAKTIDDDRVLLCGTNAKPLCRHYALKD-----GDYVVEKEVEGRLGCPFPDHNST 165
QY 175 ALFADGKLYSATVTFPLAIDANITYSLGESPTLFTVKKDSKWLKRPYFQAVDGYDYTYF 234
DB 166 AYSYGQYLSATVAOFSGTDPLTYRG-----PLRTERSDLKQLANPNFVNMEINDPFF 220
QY 225 FPRELAVENYTKGVFFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNSVGDGSHFYFN 294
DB 221 FPRELAVENYTKGVFFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNSVGDGSHFYFN 279
QY 295 ILQAVTDVIRIN-GRDV--VIALTFSTPYNSIPGSAVCAVDMLADIASVTGTGFKQKSPD 350

```

```

DB 280 EIQSTSDIIEGNYGGQVEKLIYVFTTPVNSIGCSAVCAFPKSKSLIESDGPKEGFTNN 339
QY 351 STWTFVDEPDERVPPKPPGCCAGSSSLERYATSNFPPDDTLNFKTPELMDAAPSINRPW 410
DB 340 SNMLAVPSLKAVEPFRPGQVND-----SRTLDPVSNVFPVSRMLDEAVPAFTRP1 391
QY 411 FLRTWVRVRLTKIAYD---TAAGPYONHTVVFLSEKGIILKFLARIGNSGFLN----- 461
DB 392 LIRISLQYRFTKIAVDQOVRTPDG--KAYDVLFIGDIDDKVIAL---NSASFDSQDTV 445
QY 462 DSLFLEEMSVNYSKESQSYGVDEKRIWGMQDLBASSLVAFSTCVIKYPLGRG--ERHG 519
DB 446 DSVVIEELQVLP-----PQVYKLYVVRMGSDSKLVVSGDELLAKLHCGSDKIT 499
QY 520 KCKTICIASRDPYCGWIKEGACSHL-SPN---SRLTFQDIERNGTDLGDCNSFWAL 575
DB 500 NCRF-CVSLQDPYCANVDNELKCTAVGSDWSAGKRRFQNLISGHHKACGRPRQTEIV- 557
QY 576 NHHSSLLPSTTSDSTA-----QEGYESRGMDLDMKHLDSPSTDPDLGAVSSHNQ 628
DB 558 ---ASPVPTQPTKSSGDPVHSIQAEPEPE---IDNEIVIGVDSDSNVAPNTLAIINHA 610
QY 629 DKKGVRBSYKKGHDQVLPV---TL-LAI-----AVILAFVMAVFS 666
DB 611 GSK-----LPSSQEKPLPYTAETLITLIALVTSCLGALVVGFTSGFLFS 652

```

```

RESULT 10
PCT-US94-10151A-58
Sequence 58, Application PC/TUS9410151A
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESSES:
ADDRESS: FLEHR HOBBACH TEST ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10151A
FILING DATE: 13-SEP-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-10151A-58

Query Match      17.8% Score 969.5; DB 5; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LYFTLLHRAAGFPPDSSEPTISHSNNTKQYVFGHKGKNTTOR-----HRLDQMI 61

```

```

Db      11 LLMVALH-AAAMVNDVSP-----KMYVQF-----GEERVOFPLGNESHKHEKLL 54
QY      62 MIMNGGLYIARADHIYTVID--TSHTEEIYCSKLLTWKROADVDTCKMKKHDECHN 119
Db      55 EKDHNSLVGANNIYINISRLDTEFTEQ-----RIEMHSSAHNELCTLKKSXDDCN 109
QY      120 FIKVLKKNDDALFVCGTNAFNPSCNRYMDTLEPPGD-----EFGMARCPYDAKHANV 174
Db      110 YIRVLAKIDDDVLLIGTNAYKPLGRHYALKO-----GDYVVEKEVEGRGLCPFPDDHNS 165
QY      175 ALFADGKYSATVDTPLAIDAVIYRSLGSPILRTYKHDSKYLKEEYFVQADYGDYIYF 234
Db      166 AIYSEGLYSATVADPSGTDPLIYRG-----PLKTERSDLKQINAMFNVTNYEYDFLTF 220
QY      235 FPREIYEVNTMGKVPFPPVAOVCKNDMGSGORVLEKOWTSPFLKARLNCVSGDSHFYFN 294
Db      221 FPREIYEVNTMGKVPFPPVAOVCKNDMGSGORVLEKOWTSPFLKARLNCVSGDSHFYFN 279
QY      295 ILQAVTVIRIN-GRDV---VLAFTSTPYNSIPGSANCAVMDLILASVFTGPRKQKSPD 350
Db      280 BIQSTSDIIEGNVGGVEKLIYGVFTTPVNSIGSAYCAFAFMKSLIESFDGPFKEQETWN 339
QY      351 STWTPYDPERVPRKPRGCCAGSSSLERVATSNFPDDTLNFKTHPLMEAVPSIENRW 410
Db      340 SNMLAVPSLKVPRPRGQCVND-----SRTLPVSVNPFVKSHTLMDEAVPAFTREI 391
QY      411 FLATWRYRLTKLAVD-----TAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFLN----- 461
Db      392 LIRISIQRYRTKLAIVQOVRTPDG--KAYDVLFITGDDGKVIKAL-----NSASFDSQDVT 445
QY      462 DSLFLEMSVYNSEKSYDGEDKRIHQMDLRASSLLVYAFSTCYIKPPLGRC--ERBG 519
Db      446 DSVVIEELQVLP-----PGVPYKNLYVMMDGDDSKLVVSDDELIAIKLHRCGSDKIT 499
QY      520 KCKKTCLASDPPYCGMIKEGGAASHL-SPN---SRLEFQDIERGNTDGLGCHNSFVL 575
Db      500 NCRE-CVSIODPPYCAMDNVELKCTAGSSPWSAGKRFFIQNISLGHKACGGRPQTEIV- 557
QY      576 NGHSSSLPSTTSDSTA-----OEGYBSRGMDLWKHLLDSPSTDPGLAVSSHNQ 628
Db      558 ---ASVYVQPTTKSSGDPVSHIHQAFEPPE---IDNELVIGVDSBNV.PNTLAEINHA 610
QY      629 DKKGVIRESYLKGHDLVPTV---TL-LAI-----AVIIAFVWGAIVES 666
Db      611 GSK-----LPSSQEKLPYVTAETLLIATVTSCLAGLVGFLSGFLFS 652

```

```

RESULT 11
US-08-121-713D-60
; Sequence 60, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D

```

```

; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 650 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-121-713D-60

```

```

Query Match      16.1% Score 880; DB 1; Length 650;
Best Local Similarity 35.8% Pred No. 1.3e-74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

```

```

QY      114 KDECHNFTKVLKKNDDALFVCGTNAFNPSCNRYMD-----TLEPPGDHFGMARCPYNA 169
Db      1 EDDCCNYIRIMVVPSPGRFLFVCGTNSFRPMCNVTYIISDSNYTLBA---TKNGQAVCPYDP 57
QY      170 KHANVALPADGKLYATVDTPLAIDAVIYRSLGSEPTLRTVGHDSKYLKEEYFVQADYDG 229
Db      58 RHNSTSVLADNDELVSGTADFSGDPILYRE-----PLQTEQYDLSLNAFPVSSFTQG 112
QY      230 DYIYFPPREIAYEVNTMGKVPFPPVAOVCKNDMGSGORVLEKOWTSPFLKARLNCVSGDS 289
Db      113 DSVYVFFRRTAVFPLNGCAIYSRVARVCKMKGGPR-FRNRWTSFLSRINCSPIDGY 171
QY      290 HEYFNILQAVTVIR---INGRDVVLAFTSTPYNSIPGSANCAVMDLILASVFTGPRERE 345
Db      172 PFYFMEIQASNLIVEQYSGMSKLIYGVFTTPVNSIPGSANCAVMDLILASVFTGPRERE 231
QY      346 QKSPDSTWTPYDPERVPRKPRGCCAGSSSLERVATSNFPDDTLNFKTHPLMEAVPSI 405
Db      232 QGINSNMLPVNNAKVPPRPPSG-----HNSRALPDPLNFKTHSLMDENVPAF 283
QY      406 FNRPFILRTMVRVRLTKLAVD-----TAAGPYQNHVTVFLGSEKGIILKFL-ARIGNSGFL 460
Db      284 FSPILVIRSTIYRFQIAVDAQIKTPG--KTYDVIFVGTGDKLIKSYNAESADSADK 341
QY      461 NDSLFEEMSVYNSEKSYDGEDKRIH-QMOLDRASSSLY-----VAFSTVIRKPLG 513
Db      342 VTSVVEIEIDVLTKS-----EPINLEIVRTMQDQPKDGSYDDGKLIIVTDSQVVAIQH 397
QY      514 RC--ERHGCKKCTCIASRDPYCGMIKEGGAASHLSPN---SRLEFQDIERGNTDGLGDC 568
Db      398 RCHNDKITSCSE-CVALDPPYCAMDKIAGCRSHGAPRMLSENYFFQNVATGQ----- 449
QY      569 HNSFVALNGHSSSLPSTTSDSTAQOEGYBSRGMDLWKHLLDSPSTDPGLAVSSHNQ 628
Db      450 -----HAACPSGKINSKDNAMAGEQKFRNDM-----DLIDS-----RQ 483
QY      629 DKKGVIRESYLKGHDLVPTVTLAIVIIAFVWGAIVES 666
Db      484 SKDQEIINDIKNFEIINAGYVETIVVAVLGSLFS 521

```

```

RESULT 12
US-08-835-268-60
; Sequence 60, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy

```



```

Db      172 PFYNEIISASNLVEGQYGSMSKLIYGVFTSPNSIGSAVCAFALQDIADTFEGQFKE 231
Qy      346 QKSPDSTWTPVDERVPRKPPGCCAGSSSLERVATSNFPDDTLNFTKTHPLMDAVPSI 405
Db      232 QTGINSNMLPVNNAKVPDRPGSC-----HNSRALPPTLNFITKTHSLMDENVPF 283
Qy      406 ENRPFLRTWRYRLTKIAVD-----TAAGPYONHTVFLGSEKGIILKFL-ARIGNSGFL 460
Db      284 FSQPLVIRITSTYRFTQIAVDAQITPGG--KTYVIVIVGDHGIKISVNAESADSDX 341
Qy      461 NDLFLFEMSVYNSKCYDGVDEKRIK-GMQLDRASSSLY-----VAFSTCVIKVPLG 513
Db      342 VTSVIVIEIDVLTGS-----EPINLEIVRTWQYDQPKDGSYDDGKLIIVTDSQVVAIQLH 397
Qy      514 RC-ERHGKCKKTCCLASDPYCGWIKEGAC-SHLSFN--SLTFEODIERGNTGLDGC 568
Db      398 RCHNDKITSCE-CVALDPPYCAMDKIAGKCRSHGAPRWLENNYVQNAVATQ-----449
Qy      569 HNSFVALNGHSSSLPSTTSDSTAOEGYESRGMLDWKHLSDPSDTPDLGAVSSHNO 628
Db      450 -----HAACPSGKINSKDNANGEQKGFNDK-----DLDS-----RRQ 483
Qy      629 DKGVIRESYLKGHDQVLTLLAIAVILAFVWGAVFS 666
Db      484 SKDQELIIDNIDKNFEDIIAQTVEITLVAVLAGSIFS 521

RESULT 14
US-08-833-391-60
; Sequence 60, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 650 amino acids
; TYPR: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

US-08-833-391-60
Query Match      16.1%; Score 880; DB 3; Length 650;
Best Local Similarity 35.8%; Pred. No. 1,3e-74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

Qy      114 KDEHNFIKYLKKNDDALFVCGNAFNPSCRNYKMD-----TLPPDEFSGMARCFYDA 169
Db      1 EDDQNTIRIMWVDSPPGLFVCGINSFRPMCNITYIIDSNTTLEA-----TKNGQAVCPDP 57
Qy      170 KHANVALFADGKLVSATVDTFLAIDAVYLSLGSFTLRIVKDSKWLKEPFIQAVDYG 229
Db      58 RHNSTVLADNELISGVTVADPSGSDPIIYRE-----PLQTEQYDSLINAFNFSSTFQG 112
Qy      230 DYIIFPFRRLAEVNTWKVFPFVAVQYCKNDMGSGQVLEKQMTSPFKARLNCSPVDS 289
Db      113 DVFVFFRETVAFEPINCKALYSRVARCKWDKGGPHN-FENRMTSPKSLNCSIPDGY 171
Qy      290 HVENIILQAVTVDR-----INGRDVLTAFSTPYNSIPGSAVCAVMDLIDAVFTGRPKE 345
Db      172 PFYNEIISASNLVEGQYGSMSKLIYGVFTSPNSIGSAVCAFALQDIADTFEGQFKE 231
Qy      346 QKSPDSTWTPVDERVPRKPPGCCAGSSSLERVATSNFPDDTLNFTKTHPLMDAVPSI 405
Db      232 QTGINSNMLPVNNAKVPDRPGSC-----HNSRALPPTLNFITKTHSLMDENVPF 283
Qy      406 ENRPFLRTWRYRLTKIAVD-----TAAGPYONHTVFLGSEKGIILKFL-ARIGNSGFL 460
Db      284 FSQPLVIRITSTYRFTQIAVDAQITPGG--KTYVIVIVGDHGIKISVNAESADSDX 341
Qy      461 NDLFLFEMSVYNSKCYDGVDEKRIK-GMQLDRASSSLY-----VAFSTCVIKVPLG 513
Db      342 VTSVIVIEIDVLTGS-----EPINLEIVRTWQYDQPKDGSYDDGKLIIVTDSQVVAIQLH 397
Qy      514 RC-ERHGKCKKTCCLASDPYCGWIKEGAC-SHLSFN--SLTFEODIERGNTGLDGC 568
Db      398 RCHNDKITSCE-CVALDPPYCAMDKIAGKCRSHGAPRWLENNYVQNAVATQ-----449
Qy      569 HNSFVALNGHSSSLPSTTSDSTAOEGYESRGMLDWKHLSDPSDTPDLGAVSSHNO 628
Db      450 -----HAACPSGKINSKDNANGEQKGFNDK-----DLDS-----RRQ 483
Qy      629 DKGVIRESYLKGHDQVLTLLAIAVILAFVWGAVFS 666
Db      484 SKDQELIIDNIDKNFEDIIAQTVEITLVAVLAGSIFS 521

RESULT 15
US-09-060-610-60
; Sequence 60, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610

```

FILED DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/835,268
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-610-60

Query Match 16.1%; Score 880; DB 4; Length 650;
Best Local Similarity 35.8%; Pred. No. 1.3e-74;
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;

QY 114 KDECNFIKVLKKNDALFYCGTNAFNPSCRNKMD---TLPEFGDEFGMARCPYDA 169
DB 1 EDDCGNYIRIMVPSGRFLVCGTNSFRPMCTYIISDNYTLFA--TKNGQAVCPYDP 57

QY 170 KHANVALPADGKLYATYTDFAIDAVIYRSLGESPTLRIVYHDSKMLKEPFIQAVDYG 229
DB 58 RHNSSTVALADNELYSGTVADFGSDPIIYRE---PLQTEQYDSLAINAPFVSSFTQG 112

QY 230 DYVFFPEEIAEYNTMGVVFPRVAQVQNDMGSGQRYLEKQWTSFLKARLNCVPGDS 289
DB 113 DVIYFFPEETAVEFINGCAIYSRVAVCKMDKGGPHR-FRRKWTSLKSLNCSIPBDY 171

QY 290 HEYFNILQAVMDVIR---INGRDVLAFTSTPYNISIPGSAVCAYMDLIAVFTGRFEK 345
DB 172 PFYFNEIQSASLVGQYGSMSKLIYGVNTPSNSIPGSAVCAPALQIDATFEGQFEK 231

QY 346 QKSPDSTWTPVDERVPRPCCCAAGSSSLERYATSNFPPDITNFIKTHPLMDEAVPSI 405
DB 232 QTGINSNMLPVNNAKVPPDRPGSC-----HNSRALPDPTLNFIKTHSLMDENVPAF 283

QY 406 ENRPWFLETRMAYRRLTKIADV---TAAGPYQNHVTVFLGSEKGIILKPL-ARIGNSGFL 460
DB 284 FSQPIIVRTSTIYRFTQIADVDAQIKTPGG-KTYDVIYVGTDHGKLIKSVNAESADSADK 341

QY 461 NDSLPLEMSVYNSEKSYDGYDEKRIK-GMQLDRASSSLY-----VAFSTCVIKVPLG 513
DB 342 VRSVVIIEIDVLTKS-----EPTRNLEIVRTMOWDQKDSYDDGKLIIVTDSQVVAIQH 397

QY 514 RC-ERHGKCKTCTIASRPDYGMIKEGAC-SHLSPN--SRUTEQDIERGNTGSLGDC 568
DB 398 RCHNDKITSCSE-CVALQDPYCAMDKIACKRSHGAPRWLEENYFYQNAVATGQ----- 449

QY 569 HNSFVALNGHSSSLDPTSTDSATQEGYESRGMIDMKHLDSPDSTPLGAVSSHNQ 628
DB 450 -----HAACPSGKINSKDNAGEQKGFNDM-----DLIDS-----RRQ 483

QY 629 DKKGVIRESYLKGDQIVPVTLLAIVLIAVFMGAVFS 666
DB 484 SKDQETIINDIKNFEDIINQYVETLVAVLAGSIFS 521

Search completed: October 23, 2003, 17:10:59
Job time: 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 23, 2003, 17:09:07 ; Search time 26 Seconds
(without alignments)
3809.759 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450

Sequence: 1 MRSEALLLYTLHFAGAGF.....PPKSPAPLSTSMKNDPACT 1030

Scoring table: BLOSUM62

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Result No. Score Query Match Length DB ID Description

1	971.5	17.8	730	2	JH0798	fasciclin IV precu
2	875	16.1	656	2	B49423	semaphorin I - fru
3	852.5	15.6	771	2	D49423	semaphorin III precu
4	842.5	15.5	711	2	A49423	semaphorin I precu
5	842.5	15.5	772	2	I48747	semaphorin D - mou
6	839.5	15.4	772	2	A49069	collapsin - chicke
7	830	15.2	1074	2	JC5928	semaphorin F precu
8	826	15.2	749	2	G01856	semaphorin V - hum
9	803	14.7	748	2	I48744	semaphorin A - mou
10	793.5	14.6	666	2	I58169	semaphorin III - m
11	789	14.5	712	2	I27165	hypothetical prote
12	779.5	14.3	724	2	C49423	semaphorin II prec
13	768	14.1	753	2	G02173	semaphorin III fam
14	737	13.5	751	2	I48748	semaphorin E - mou
15	692	12.7	834	2	S66498	M-sema F protein P
16	656	12.0	782	2	I48746	semaphorin C - mou
17	641.5	11.8	760	2	I48745	semaphorin B - mou
18	356.5	6.5	653	2	T03102	semaphorin homolog
19	322.5	5.9	676	2	T03102	hypothetical prote
20	238	4.4	1945	2	T13937	plexin A - fruit f
21	217	4.0	403	2	E42521	A39R protein - vac
22	208	3.8	441	2	S29921	hypothetical prote
23	199.5	3.7	1884	2	UC4975	plexin 2 precursor
24	182	3.3	1905	2	TS1553	plexin - African c
25	163.5	3.0	3968	2	A44265	trithorax homolog
26	158	2.9	2051	2	T13164	plexin B - fruit f
27	155.5	2.9	1894	2	UC4980	plexin 1 precursor
28	155	2.8	1375	1	UC5148	hepatocyte growth
29	155	2.8	2352	2	T30201	Notch homolog prot

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

30	153	2.8	625	2	S48941	regulatory protein
31	147.5	2.7	3507	2	T34513	hypothetical prote
32	147	2.7	1425	2	T30811	hepatocyte growth
33	146.5	2.6	295	2	UC1775	salivary protein - v
34	140	2.6	1390	1	TVHOME	hepatocyte growth
35	139.5	2.6	3869	2	A48205	All-1 protein +GTE
36	138.5	2.5	1065	2	S19482	hypothetical prote
37	138	2.5	728	2	S48569	hypothetical zinc
38	137	2.5	867	2	T41308	hypothetical poly
39	137	2.5	2492	1	C44213	ubiquitous RPR mot
40	136.5	2.5	1347	2	T02214	oxysterol-binding
41	136	2.5	1310	2	T40135	notch protein homo
42	136	2.5	2531	2	S18188	RM1 protein - mous
43	133.5	2.4	1829	2	T14280	nascent polypeptid
44	133.5	2.4	2187	2	T30826	notch-1 protein -
45	133.5	2.4	2531	2	A46019	

ALIGNMENTS

RESULT 1

JH0798
fasciclin IV precursor - American bird grasshopper
C:Species: Schistocerca americana (American bird grasshopper)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: JH0798
R:Kolodkin, A.; Kachtes, D.J.; O'Connor, T.P.; Patel, N.H.; Admon, A.; Bentley, D.; Good
Neuron 9, 831-845, 1992
A>Title: Fasciclin IV: Sequence, expression, and function during growth cone guidance in
A:Reference number: JH0798; MUID:93040225; PMID:1418998
A:Accession: JH0798
A:Molecule type: mRNA
A:Residues: 1-730 <KOL>
A:Cross-references: GB:L00709; NID:G160844; PID:G160845
A:Experimental source: embryo
C:Comment: This protein plays a role in growth cone guidance in the developing central n
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-730/Product: Fasciclin IV #status predicted <EXT>
F:23-627/Domain: extracellular #status predicted <EXT>
F:628-652/Domain: transmembrane #status predicted <TM>
F:653-730/Domain: intracellular #status predicted <INT>
F:44,71,163,267,360,539/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.8% Score 971.5; DB 2; Length 730;

Best Local Similarity 34.4% Pred. No. 5.4e-57; Indels 117; Gaps 29;

Matches 244; Conservative 124; Mismatches 224;

8 LYFTLLHFAGAGFPEDSEPISTHGNVTKQPVFVGHKPGRTTQR-----HRDDIQMI 61

11 LIIWALH--AAAVNDVDP-----KMYVQF-----GEERVGRFLGNESHKDFKLL 54

62 MIMNGTLYIARADHIYTDID--TSHTTEIYCSKLTWKSROADVDTCRMRKHDECGN 119

55 EKDNHSLVGNANIVYINSLRLDTEFTQ-----RIEWHSSGAHRELCTLKESSEDCON 109

120 FTKVLKNDALFPCGNAPNPSCRNKMOTLFEFGD-----EESGMARCPYDAKANAV 174

110 YKRVLAKTDDDDVLLCGTNAYKPLCRHYALKD---GDVVEKYEKGRGLCFDPDPHNT 165

175 ALFADGKLYSATVTPLAIDAVIYRSLGESPTLRIVKHDSKWLKEYPVQAVDGYIYF 234

166 AYSRGQLYSALVADPSGIDPLIYG-----PLRTSRDLKQJNAPNFTMEYNDPIFF 220

235 PREFAVEYNTMGKVFPPVAVOYCKNDMGSGQVLEKQMTSLKALNCSVGDSEFFYN 294

221 FRETAVEYINCGKALYSVARVCKDKGPPQFGR-WTSFLSKLNSVSGDYFFYN 279

295 IIOAVTDVIRIN-GRDV---VLATSTPYNSIPGSAVCAVYMDIISVTFGRFKEQKSPD 350

280 EIQSTSDIIEGNGVGOVEKRLIYGVFTTPVNSIGGSVCAVCAFSKMSILSESDGFFKBEETWN 339

```

QY 351 TWTVPBERPKRPGCGCAAGSSLLERYATSENEPDTLTNEIKTHPLMDEVPISIFRPN 410
Db 340 SMTLAVPSLKPBERPPGQCVND-----SRTLPDVSAFVKSHLTLMDEAVPAPFTIRPI 391
QY 411 FLRTWVRRLTLKIAVD-----TAAGPYQNTVTVFLGSEGLILKELPARIQNSGFUN----- 461
Db 392 LIRLSIQRFRIKIAVDQQVIRTPDG--KAYDVLFIGTDDGKVIKRL---NNSPSSDVT 445
QY 462 DSLFLEMSVYNSSEKSYDGYVEDKRIMGQIDRASSLIYAFSTCVIKVPYGRG--ERHG 519
Db 446 DSVATLEQLVLP-----PGVPVKNLVYVRMDQDSDSKLVVSDDEIILAIKLRGSPKIT 499
QY 520 KCKKTCIASRDPYCGMIKEGACSHL-SPN---SRLPEDQIIEBNTDGLGDCNSVAL 575
Db 500 NCRB-CVSLQDPYCGMNDVVELKCTAVGSPDMSAGRRFQIOMISGHEKACGGRQDETIV- 557
QY 576 NGHSSLLPSTTSDSTA-----QEGYESRGKMLDMKHLIDSPDSTDELGAVSSHQ 628
Db 558 ---ASVPPTOPTKSSGDVPVSHIQAEFEPE--IDNEIIVIGVDDSVAVINTLAEINHA 610
QY 629 DKGVIRESYLKGHDQLVV-----TL-LAI-----AVILAFWVGAVPS 666
Db 611 GSK-----LPSGQKLPITYAETLTIAITYSCGLALVVGFISSGLFS 652

```

QY	114	KDECHNFKVLKKNDDALFVCGTNAFNPSCNRYMD----	TLEPGEFEGSMARCPDA	169
		
DB	1	EDDCQGVYIRIMVPSGRFLFCVGINSFRMCNTYIISDSNYTLA----	TKGQGAQCPYDP	57
		
QY	170	KRANVALFPADGLKYSATVDELDAIVYRSIGSEPTLRVGHDSKMLKEPFYQAVDYG		229
		
DB	58	RHNSTSVLADNELYSGTADFGSGDPITYRE----	PLQTEYDYSLSLAPYFVSSFTQG	112
		
QY	230	DYIYFFFEIAYEVNMGKVFEPYVAQVCXNMGSORULEKQWSTFLKARINCSVPDS		289
		
DB	113	DYVYFFFEETAYEFLNCGALITSRYAVCKMKKGPHR-FRNRMTSFLKSLRMCSIPDGY		171
		
QY	290	HEFYNTLQAVTVIR----	INGRDVLAFTSPYNISIPESAVCAVYMDLDIASVFGGREKE	345
		
DB	172	PFYFNELIQSASNLVEGOYSMSKLIYGFNTPNPSNIPSAVCAFAFLQIADTFEGQFPE		231
		
QY	346	QKSPDSTWMPVPEREPRKRPQCGAGSSLEBYANSNEPDDTLNFIKTHPLMDAAPS		405
		
DB	232	QIGINSNMFLPWNNAKVPDRPGSC-----	HNSRALPDPLTNFIKTHSLMDANVPAF	283
		
QY	406	FNRPFELRTMVRVRLTKIAYD----	TAAQPYONHTVVFGLSEKGIILKFL-ARIGNSGFL	460
		
DB	284	FSGPIIVRTSTIYRFQIAVADAQIKTPGG-KTYDYIVFGTDHGKLIKSVNAESADSDAK		341
		
QY	461	NSLFLFENSVNSSEKCSIDGVEDKRIY-GMQLDRASSSLY-----	VAFSTCVIKVPLG	513
		

Db 342 VTSVIEIEDVLTGS-----EPIRNIETRMQYDDPKDGSYDGLKLLIVTDSQVALLQLH 397

Qy 514 RC-ERNKCKKCTCIASRDYCGWIKKGAC-SHLSPN-SHLTFQDIERGNTDGLGDC 568

Db 398 RCHMDKITSCEE-CVALQDPYCAMDKXGKCRSHAPMLEENIYQVAVATQ----- 449

Qy 569 HNSFVALNGHSSLLPFTTSDTAQGEYESRGGKLMWKHLIDSPDSTDFLGAVSHHQ 628

Db 450 -----HAACPGKINSKDANNGEGKFRNDV-----DLIDS-----RRQ 483

Qy 629 DKGYIRESYLGHD-----QLVEPTLLAIVLAFVNGAVS 666

Db 484 SKDGEIINIDKNNEGQTSADIINAAQYVTEVLMAVLAGSIFS 527

Db	429	YQFTQIVVRVDADEQYDMFPGTGVGLKVASLPKFTWYDLBEVLLLEMTVFR----	484
Qy	418	YRLTKTAVDTAAGYQNNHYVFGSEKGIILKTLKAIKNGSFLNDSLFLPEMSVYSEK	477
Db	370	QGSVPYPRFGTGP-SKTFGGPSTKDLDDVITFPAHSHAMVYVPPMNNRPVITKDVA	428
Qy	358	DERVPPRPQCCAGSSSLERLYATSNFPDITNFIKTHPLMEAVPSIFNRWFLTWR	417
Db	311	VFLMNRKDPKNPVVYGVFTTSSNIFKSGAVCMSSDVRVFLGFAHHDGPNYGVY-	369
Qy	302	VIRINGRD----VYLATFSTPNYSIPGSAVCAVMDLIDIAVFGPRKEQKSPDSTW	357
Db	257	SGKATHARIGQLCKDNFGG-HRSLVNMKWTTFELKALICVPGPNGTDFR----	310
Qy	246	MGRVYFPRVAQVCKNDMGSSQRYLEKQWTSFLKARINCVPG---DSHFYNILOAVTD	301
Db	197	RDFALFRITLGHHPRIITEQHDSRMLNDPKFISAHILISESDNEDDDKVYFFPRENAIDGH	256
Qy	193	IDAVIYRSLGESPRLRYVHGDSKWLKEPFVQA-----VDGYIYFEPREIAVEYNT	245
Db	137	FHPICITYIEIGHPEPDNI FKLENSHFENRGKSPYDPKLTLSLIDGELYSGTAADEFG	196
Qy	140	ENPSCR----NYKMDLTFPQDEF--SGMACPPDAKHAANALPADGLYSATVTDFA	192
Db	82	L-----VNIKDPQKILWVPYSYTRRDECKWAGKDILKECNFIKVLKAYQTHLYACGTA	136
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81
Qy	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Db	208	Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Qy	Query Match	15.6%; Score 852.5; DB 2; Length 771;	
Db	Best Local Similarity	32.1%; Pred. No. 5.5e-49;	
Qy	Matches	208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Db	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81
Qy	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Db	208	Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Qy	Query Match	15.6%; Score 852.5; DB 2; Length 771;	
Db	Best Local Similarity	32.1%; Pred. No. 5.5e-49;	
Qy	Matches	208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Db	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81
Qy	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Db	208	Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Qy	Query Match	15.6%; Score 852.5; DB 2; Length 771;	
Db	Best Local Similarity	32.1%; Pred. No. 5.5e-49;	
Qy	Matches	208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Db	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81
Qy	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Db	208	Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Qy	Query Match	15.6%; Score 852.5; DB 2; Length 771;	
Db	Best Local Similarity	32.1%; Pred. No. 5.5e-49;	
Qy	Matches	208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Db	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81
Qy	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Db	208	Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Qy	Query Match	15.6%; Score 852.5; DB 2; Length 771;	
Db	Best Local Similarity	32.1%; Pred. No. 5.5e-49;	
Qy	Matches	208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Db	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81
Qy	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Db	208	Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Qy	Query Match	15.6%; Score 852.5; DB 2; Length 771;	
Db	Best Local Similarity	32.1%; Pred. No. 5.5e-49;	
Qy	Matches	208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Db	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81
Qy	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Db	208	Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Qy	Query Match	15.6%; Score 852.5; DB 2; Length 771;	
Db	Best Local Similarity	32.1%; Pred. No. 5.5e-49;	
Qy	Matches	208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;	
Db	44	HKPGNTQRHLDLQIMINMGT-----LYIARDHIYTVD	80
Qy	81	IDTSHTEIYCSKLTWKSROADVDTCRKKGK-KDECHNFIKVLTKXNDALFVCGTNA	139
Db	22	YONGKNNVPRLLSYKEMLESNNVITFNGLANSSYHFFLDEESRLYVGAKDHFSD	81

```

Db      485 -----EFTAIASNELSTKQOOLYIGTAGVAQLPHHCDDLYGKACACCCARPYCAM-- 537
QY      538 EGACGSHLSFNS-RLTEQDIERGNTDGLDPC---HNSFVALNGHS--SSLIPSTTSD 590
Db      538 DGSACSRFFPAKRTTRQDRIANG--DPLTHCSDLHHNDH---HGHSPERIIVGVENS 592
QY      591 STAOEGESRGKMDMKHLSDPSTPLGAVSHNHQDKKGVATRES 637
Db      593 TFECSKRSQRAIVYQFORNERKEDEI-RVDHIIIRTDGILLRS 638

```

RESULT 4

```

A49423
semaphorin I precursor - beetle (Tribolium confusum)
C/Species: Tribolium confusum
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C/Accession: A49423
R/Klodkin, A.L.; Matthes, D.J.; Goodman, C.S.
Cell 75, 1389-1399, 1993
A/Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
A/Reference number: A49423; MUID:94094332; PMID:8269517
A/Accession: A49423
A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A/Molecule type: mRNA
A/Residues: 1-711 <KOI>
A/Cross-references: GB:I26080

```

```

Query Match      15.5%; Score 845.5; DB 2; Length 711;
Best Local Similarity 28.9%; Pred. No. 1.4e-48;
Matches 230; Conservative 130; Mismatches 270; Indels 165; Gaps 30;

```

```

QY      12 LILHAGAGFPDSEPI-----SISHGNTKQYRVFVGHKPKRNTTQHRLDIQIMIMN- 65
Db      12 LIALCHAMMPDSSSKLTHFKSVESKSEFT-----GNATFPDH-----FVILNQ 54
QY      66 --GTLTYAARPHITVDI--DTSHTBEIYCSKLLTWKSRQAVDTCRMKSKHKECHNPIK 122
Db      55 DETSILVGGRRVYNLSIFDLSEK---GGRIMPSDDAHGQICILKGTDDDCQNIYR 110
QY      123 VLLKKNDALFVCGTNAFNSCRNYKMDLEPFGD-EFSGARCPYDAKMANVALFADGK 181
Db      111 ILTSEPEKLVICGTNSYKPLCRTPAFKEGKYLVEKEVEGIGLCPRYNEHNSISVYNGQ 170
QY      182 LYSATVDFLAIDAVIYRSLGESPTLRTVKNDSKWLKEPFYVQAVDYDYIYFFREIAY 241
Db      171 LFSATVADFGSGDPLIYRE-----PORTELSDLKQINAPFNVSVAVDYIYFFRYRTAV 225
QY      242 EYNTMGKVPFVAVOVCKNDMGSGQVLEKQWTSFLKARLNCVSGDSHFYFNILQAVTD 301
Db      226 EYMGCKVITYSRVAVKCKDQKGRHQRDR-WTSFLKARLNCSTIGBEYFFYFDELQSTSD 284
QY      302 VI--RINGRD--VVLATFSTPYNSIPGSAYCAVMDIASVFTGRFKEKQSPDSTWTPV 357
Db      285 IVEGYNVSDDEKLIIGLITTPNALIGSALICAYQVADILRVEGSGFKQETINSWMLVP 344
QY      368 DERVYKPRPGCCAGSSSERATSNFPPDPLNFKITPLMDPAVPSIFNKPWELRTVVR 417
Db      345 QNLVEPRPGQCVDRSRI-----LPDKVNFIKTSHMED-VPALEFGKVLVRLVSLQ 395
QY      418 YRLTIADVTAAGPYQNH--TVVPLGSEKGIILKFLARIGNSGFLNDSLFLIEMSVYSE 475
Db      366 YRFTLITVDPOVKTTNNQYLVDLYIGTDDGKYLK----- 429
QY      476 KQSYDVEDDKRLMGQDLRASSSLVYAFSTCY-----IKV----- 511
Db      430 -----AVNIPKHAHALLYRKYRTSVHPGARPKYKLIAGYKVVVVGKDEIR 478
QY      512 ---LGRCEHKGCKTCTIASRDPYGMKKEGACSHLSFNSLTF--EQDIENG----- 561
Db      479 IANLHNCASKTRC-KDQVELQDPCAMDAKQNLCVSITLVISYRFLIDIVNAGDNNKWS 537
QY      562 --TDGLGCHNSFVALNGHSSILPSTTSDTAQEGYESRGKMDMKHLSDPSTDPL 619

```

```

Db      538 PQDKDKTVLKNK-----PSEVENBEIT-----NSIDEKDL-----DSSDPL 572
QY      620 GAVSHNHQDKKGVATRESITKG---HDQLYPV--LLATAVILAFWGA---VFSGITVY 671
Db      573 IKTGLDDSDCPDV--SENSIGCAVRQOLVITYAGTLHYVVVAVSIVGLFSWYLSGLSVF 631
QY      672 CV--CDHRKADVAVQKKE--KELTSPRSGMS--SVTKLSGLFGDTCQSKPKREALTPL 726
Db      632 AKHSDSQPEAPFIOHNLHERLSANQGYLTPRANKAVNLVYKXSSSTPRPKMDLV 691
QY      727 MHNGKLTAPGNTAKM 741
Db      692 SKDLNIAISDTLOKI 706

```

RESULT 5

```

148747
semaphorin D - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C/Accession: I48747
R/Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A/Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates c
A/Reference number: I48744; MUID:95267431; PMID:7748561
A/Accession: I48747
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-772 <RES>
A/Cross-references: EMBL:X85993; NID:9854329; PID:CAA59985.1; PID:9854330
A/Genetics:
A/Gene: semD
C/Superfamily: semaphorin

```

```

Query Match      15.5%; Score 842.5; DB 2; Length 772;
Best Local Similarity 34.3%; Pred. No. 2.6e-48;
Matches 200; Conservative 85; Mismatches 225; Indels 73; Gaps 16;

```

```

QY      47 GRNTQCHRLDIQIMIMNGT-----LYIAARDHIYTDIDT 83
Db      25 GKNNVPRKLSTYKEMLESNNVITFNGIANGSSYHTFLIDBERSRLYGAKHITSFNL-- 82
QY      84 SHTBEIYCSKLLTWKSRQAVDTCRMKQH-KOECHNFIKVLKKNDDALFVCGTNAENP 142
Db      83 ---VNIDPQKIWPVSVYTRRDECKMAGKQILKCAPIKYLEAVNQHLYAACGTGAFHP 139
QY      143 SCR-----NYKMDTLEPGBEF--SGMARCPYDAKMANVALFADGKLSATVDFLAIDA 195
Db      140 ICTYIEVGHHPEDNI FKLQDSHPENGRGKSPYDKLTLASLLIDGELYSGTAADFMRGR 199
QY      196 VIYSLGESPTLRVKNDSKWLKEPFYVQV-----VDYGDYIYFFREIAYEVNIMGK 248
Db      200 AIFPTLGDHHDITREQDHSKRLNDPRFISALHPESDNPEDDKYFFPRENAIGSEHSGK 259
QY      249 VVFPFRAVQCKNDMGSGQVLEKQWTSFLKARLNCVSG---DSHFYFNILQAVTDVIR 304
Db      260 ATHARIQIQKNDRG--HRSLVNKMWTFELKARLICVGPGRGIDTHF-----DELQVFL 313
QY      305 INGRD---VVLATFSTPYNSIPGSAYCAVMDIASVFTGRFKEKQSPDSTWTPVPER 360
Db      314 MNSXDPKPIPIYVGYFTTSSNIFKGSAYCMVMSDVRRFQVLAHRODPNQPWPPY--QGR 372
QY      361 VPKRPGCCAGSSSLERATSNFPPDPLNFKITPLMDPAVPSIFNKPWELRTVVRYL 420
Db      373 VPIPRPGTCR-SKTFPGFDSTKDLPDVITIGRSHPAKYNVVFPINNRPIMIKTDVNVQF 431
QY      421 TKIADVTAAGPYQNHTVFLGSEKGIILKFLARIGNSGFLNDSLFLIEMSVYNSSEKSYD 480
Db      432 TQIVADRYADADGGQVDMFICTDVGTVLKVASVPEKETHNDLEEVLLDEEMTVFR----- 484
QY      481 GVEDKRLMGQDLDAASSSLVYAFSTCYIKVPLGCEHKGCKKCTCIASRDYCGYKEGG 540

```

Db 485 --EPTTISAMELSTKQOOLYIGSTAGVQLPLHRCDIYKAKACECLARDPYCAM--DGS 540

QY 541 ACSHLSPPNS-RLTFEODIERGNTDGLGCHNSFVALNGHSSSL 582

Db 541 SCSSRFPPKRRKTRKODIRNG--DPLTHCSDLEHDHNGHSPSL 581

RESULT 6

A49069

collapsin - chicken

C:Species: Gallus gallus (chicken)

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999

C:Accession: A49069

R:Uno, Y.; Raible, D.; Raper, J.A.

Cell 75, 217-227, 1993

A:Title: Collapsin: a protein in brain that induces the collapse and paralysis of neuron

A:Reference number: A49069; MUID:94006554; PMID:8402908

A:Accession: A49069

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-772 <UO>

A:Cross-references: GB:U02528; NID:g410078; PIDN:AAC59638.1; PID:g410079

C:Superfamily: semaphorin

Query Match 15.4%; Score 839.5; DB 2; Length 772;

Best Local Similarity 35.2%; Pred. No. 4.1e-48;

Matches 198; Conservative 87; Mismatches 220; Indels 57; Gaps 17;

QY 68 LYIARADHIYVDIDTSHTEIYCSKLTWKSROADVTCRMKGKH-KDEGNFIKVLK 126

Db 69 LVYAGADHIFRFL-----VNKEYOKIYWPVSHSRDECKVAGKDILRECANFIKVLKT 123

QY 127 KNDALFVCGTNAFNSCR-----NYKMDLFEFFGDEF--SGMARCPYDAKHANVALFAD 179

Db 124 VYQTHYACGTGAFHMCYIYEGSHREDNIRMEDSHENGKRGKSPYDPLTASLVD 183

QY 180 GKLYSATVDFLAIDAIVYRSLGESPTLTWKDSKWLKEPVVQA-----VDYGYI 232

Db 184 GELYSCTADDFMGRDPAIFRTLGHHPIRTEQHSFMDLNDPRFISAHLLPESDNPEDXI 243

QY 233 YEFRELAIEYNNMGVPRVAVQVCKNMGSGRVLEKQWTSFLKARLNCSPVG---D 288

Db 244 YEFRENALDGEHTGAKTHARIGQICKNDPFG--HRLVNMKWTFLKARLNCSPVGNGID 302

QY 289 SHFYFNLQAVTDVIRNGRD---VVLATFTPYNSIPGSAVCAVMDLIDIASVFTGRK 344

Db 303 THF-----DELQVFLMNSKDPKMPYVGVFTSSNIPKSAVCAVSMTDVRAVFLGPA 357

QY 345 EQKSPDSTWTPVDERVPKPRPCCAGSSSLERYATSNRPDPTLFIKTHPLMDAVPS 404

Db 358 HRDGPVQWVPY-QGRVVPYPRGTCR-SKTFGGFDSTKDLDEVITFARSHPMVNPVP 415

QY 405 IHRPFPRLRWVYRLLTKLAVDTAAGYQNHVAVFVGESEKILTKFLARIGNSGFNDLSI 464

Db 416 INSRPIMIKTDVYQFQIVVDVADDDGOYDMFTGTDIGYLVKAVSIPKETHLEEV 475

QY 465 FLEMSVYNSKESYDVEDKRIHQDILASSSLVYAFSTCYIKVPLGCEBHGCKKT 524

Db 476 LLEEMVFR-----EPTVISAMKISTKQOOLYIGSATGVSQLPLHRCSDYVAKACAE 526

QY 525 CIASRDYCGWIKEGACSHLSPNS--RLTFEODIERGNTDGLGCHNSFVALNGHSSSL 583

Db 527 CCLARDPYCAM--DGSRCRYFPFAKRRTRKODIRNG--DPLTHCSD---LQHNDN--- 575

QY 584 PSTTSDSTAQEGESRGMLD 605

Db 576 PSQGLEKTIYVENSSTFLE 597

C:Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000

C:Accession: J05928

R:Simmons, A.D.; Puechelet, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.

Biochem. Biophys. Res. Commun. 242, 685-691, 1998

A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candid

A:Reference number: J05928; MUID:98125554; PMID:9464278

A:Accession: J05928

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1074 <SIM>

A:Cross-references: GB:U52840; NID:g2772583; PIDN:AAC09473.1; PID:g2772584

A:Experimental source: brain

C:Comment: This protein disrupts normal brain development and leads to some of the featu

C:Genetics:

A:Gene: semaf

C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology

F:1-20/Domain: signal sequence #status predicted <SIG>

F:50-533/Domain: semaphorin #status predicted <SBM>

F:840-896/Domain: thrombospondin type 1 repeat homology <THR3>

F:971-993/Domain: transmembrane #status predicted <TM>

Query Match 15.2%; Score 830; DB 2; Length 1074;

Best Local Similarity 36.2%; Pred. No. 3e-47;

Matches 199; Conservative 90; Mismatches 211; Indels 50; Gaps 17;

QY 68 LYIARADHIYVDIDTSHTEIYCSKLTWKSROADVTCRMKGKH-KDEGNFIKVLK 127

Db 70 LVYAGADHIFRFLQ-----EDSLIOAVEMCDEKTRKACYSKSGSKSECONYIVLLV 123

QY 128 NDDALFVCGTNAFNSCRNYKMDLFEFFGDEFSGMARCPYDAKHANVALF-ADGKLYSAT 186

Db 124 GGDRLFTGCTNAFTVTCNIRSLSNALHDOISGARCPYSPQHNSTALLTAGGELYAT 183

QY 187 VTDFLAIDAVIYRSLGESPTLTWKDSKWLKEPVVQAQVADYGYIYEFRELAIEYNNM 246

Db 184 AMDPFRDPAIYRSIGILPPLRTAQYNSKMLNEPVSYSYDIGNTYFEFRENVAEHD-C 242

QY 247 GKVPVPRVAVQVCKNMGSGRVLEKQWTSFLKARLNCSPVGSHFYFNLQAVTDVIRIN 306

Db 243 GKTVFSRAARVCKNDIGG-RFLLEDTWTMKARLNCSPVGSHFYFNLQAVTDVIRIN 300

QY 307 GRDVVLAFTSTPYNSIPGSAVCAVMDLIDIASVFTGRFQKSPDSTWTPVDERVPKPR 366

Db 301 --DLIYGFITNIVNSIASAACVFNLSIAIQFSGDPFKQENSRSAWLPYPP---PNPFR 355

QY 367 GCCAGSSSLERYATSNRPDPTLFIKTHPLMDAVPSIFNRPWLRTWVRLTKLAVD 426

Db 356 QCGTVDDGLYVNLTERNL-QDAQKFLVH---EVVQPTVTVSFIEDNSRF--SHAVVD 408

QY 427 TAGPYQNHVAVFVGESEKILTKFLARIGNSGFNDLSI-FLEMSVYNSKESYDVEDKR 486

Db 409 VVQGRBALVHITLADTYGTRKAVVPLNGT---SSSCLLEIELFPERR-----RFP 458

QY 487 IMGQIDRASSLVYAFSTCVIKVPLGRCERHGCKKTCLASRDYCGWIKEGACSHLS 546

Db 459 IRSLOIHSQSVLEFVGLREHVVKIKPLKRCQFY-RFTSTCIGADPPCGMDVNMKCTSLR 517

QY 547 ENSRLT-FEODIERGNTDGLGCHNSFVALNGHSSSLST--TTSDDTAQEGESRGGM 603

Db 518 ESLSMTQWQGISIA-----CPTNLTVDGHPGVMSPTTCTHTDGSAAV-----GSC 563

QY 604 LMKKHLDSP 613

Db 564 LCRTRCDSF 573

RESULT 7

J05928

semaphorin F precursor - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999

C:Accession: G01856

R:Sekido, Y.

submitted to the EMBL Data Library, June 1995

A:Reference number: G08634

A:Accession: G01856

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-749 <RES>

A:Cross-references: EMBL:U28369, NID:G974283, P1DN:AA09138.1, PID:G974284

C:Superfamily: semaphorin

Query Match 15.2%; Score 826; DB 2; Length 749;
Best Local Similarity 34.7%; Pred. No. 3.2e-47; Indels 56; Gaps 16;
Matches 198; Conservative 88; Mismatches 228

```

QY 66 GTLYIARDHIYTVDDITSHTEIYCSKLTWKSROADVDTCKMKKH-KDECHNFIKVL 124
DB 65 GRFLVGENHVAASLNDISKR-----AKKLAMPAPVEMRECCWMAKDICTECNNFVRL 120
QY 125 LKKNDALFVCGTNAFNSCRNYKMD-----TLPEFGDESGMARCPIDAHANVAL 176
DB 121 HAYNHTLLACGTGAFHPTCAFVEVGHRAEPVLRLLP-GRIBDGKGSPPYDPHRAASV 179
QY 177 FADKLVSATVTDPLADAVIYRSIGESPILRTYKHSKMLKEPFYQAV-----DYG 229
DB 180 LVGBELYSVAADLMGRDFTIFRSIGQPSLRTBPHDSRWLNDEKFKVFWIPESNPDD 239
QY 230 DYIYFFPREIAVE-YNTWGVFPVAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSP-- 286
DB 240 DKTYFFRETAVEAAPALGRLSVSRVQICRNDVG-GRSLVNMKWTFLKARLVCVSPGV 298
QY 287 -GDSHFYFNILQAVTVDIRNGD-----VLATSTPPYNSIPGSANCAVMDLDAVFTG 341
DB 299 EGDTHF-----DQLQDVFLSSRDHRTPLLYAVFSTSSIFQSGAVCYVSMNDVRAFLG 353
QY 342 RFKEKSPDSTWTPVDEVPKRPCCAGSSSLERYATSNFPPDPTLNFIKTHPLMDEA 401
DB 354 PPAHKEGPMHGVST-QGRVPIYPRGMCP-SKTFGTSPSTKDPDVIDQFARHPLMYNS 411
QY 402 VPSIFNRPFELRTWVRYRLTKLAVDTAAGPYQNTVFLGSEKGIILKFLARIGNSGLN 461
DB 412 VLPGRPLFLQGVANFTFTQIADRVAAADGHVDFLFTGTVGLKAVISPKGSRPSA 471
QY 462 DSLFLEMSVYNSEKCSYGVEDKRIHQMDLRASSSLYVAFSTCVIVPLGRCGRHGC 521
DB 472 ECLLLEELHVFPE-----DSAAVTSMQISSKRCQLYVARSAVAQIALHRCAGRV 522
QY 522 KKTCTASRPYCGWIKEGACSHLSPNSRLTF-EODIERGNTDGDGCHNSFVALNGH-- 578
DB 523 CTBCCIARDPYCAM--DGAICTRFQPSAKARRRRQDVNRGDPSTLCGSGSSAPALLEKV 580
QY 579 -----SSSLPSTTSDSTAGGYESRGGM 603
DB 581 FGVESSAPFLECEPRSLGARVEMTFQAGV 610

```

RESULT 9

148744

semaphorin A - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: I48744

R:Fuschel, A.W.; Adams, R.H.; Betz, H.

Neuron 14, 941-948, 1995

A:Title: Murine semaphorin D/collyrin is a member of a diverse gene family and creates

A:Reference number: I48744, NID:95267431, PMID:7748561

A:Accession: I48744

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-748 <RES>

A:Cross-references: EMBL:X85990, NID:G854323, P1DN:CAA59982.1, PID:G854324

C:Genetics:

A:Gene: sema

C:Superfamily: semaphorin

Query Match 14.7%; Score 803; DB 2; Length 748;
Best Local Similarity 35.5%; Pred. No. 1.1e-45;
Matches 191; Conservative 80; Mismatches 205; Indels 62; Gaps 18;

```

QY 66 GTLYIARDHIYTVDDITSHTEIYCSKLTWKSROADVDTCKMKKH-KDECHNFIKVL 124
DB 65 GRFLVGENHVAASLNDISKR-----AKKLAMPAPVEMRECCWMAKDICTECNNFVRL 120
QY 125 LKKNDALFVCGTNAFNSCRNYKMDL-----EPFGDESGMARCPYDAKANVA 175
DB 121 HAYNHTLLACRTGAFHPTCALRWATAGTHASTGPEKLED---GKQKPYDPHRAAPS 177
QY 176 LFADGKYSATVDFPLADAVIYRSIGESPILRTYKHSKMLKEPFYQAV-----DY 228
DB 178 LVGBELYSVAADLMGRDFTIFRSIGQPSLRTBPHDSRWLNDEKFKVFWIPESNPDD 237
QY 229 GDYIYFFPREIAVE-YNTWGVFPVAVQVCKNDMGSGQVLEKQWTSFLKARLNCVSP-- 286
DB 238 DKTYFFRESAVEAAPAMGRMSRVQICRNDLG-GRSLVNMKWTFLKARLVCVSPG 296
QY 287 -GDSHFYFNILQAVTVDIRNGD-----VLATSTPPYNSIPGSANCAVMDLDAVFT 340
DB 297 VEGDTHF-----DQLQDVFLSSRDHRTPLLYAVFSTSSIFQSGAVCYVSMNDVRAFL 351
QY 341 GRFKEKSPDSTWTPVDEVPKRPCCAGSSSLERYATSNFPPDPTLNFIKTHPLMDE 400
DB 352 GFLPHKEGPTHGVST-QGRVPIYPRGMCP-SKTFGTSPSTKDPDVIDQFARHPLMYN 409
QY 401 AVPSIFNRPFELRTWVRYRLTKLAVDTAAGPYQNTVFLGSEKGIILKFLA---RIGN 456
DB 410 FVLPMGRPLFLQGVAGYFTQIADRVAAADGHVDFLFTGTVGLKAVISPKGSRPN 469
QY 457 SGFLNDLSFLEMSVYNSEKCSYGVEDKRIHQMDLRASSSLYVAFSTCVIVPLGCE 516
DB 470 S-----ECLLLEELQVFE-----DSAAVTSMQISSKRCQLYVARSAVAQIALHACT 516
QY 517 RHGCKKCTASRPYCGWIKEGACSHLSPNSRLTF-EODIERGNTDGL--GDCHNS 571
DB 517 ALGRACCECIARDPYCAM--DGAICTRFQTAARRRRQDVNRGDPSTLCGSGSSHS 572

```

RESULT 10

158169

semaphorin III - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999

C:Accession: I58169

R:Messersmith, E.K.; Leonardo, E.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S.; Ko

Neuron 14, 949-959, 1995

A:Title: Semaphorin III can function as a selective chemorepellent to pattern sensory pr

A:Reference number: I58169, NID:95267432, PMID:7748562

A:Accession: I58169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-666 <RES>

A:Cross-references: GB:I40484, NID:G703189, P1DN:AAA73934.1, PID:G703190

C:Genetics:

A:Gene: SemaIII

C:Superfamily: semaphorin

Query Match 14.6%; Score 793.5; DB 2; Length 666;
Best Local Similarity 34.1%; Pred. No. 4e-45;
Matches 188; Conservative 89; Mismatches 226; Indels 49; Gaps 15;

```

QY 116 ECHNFIKYLKKNDALFVCGTNAFNSCR-----NYKMDTEPFGDER--SGMARCPVD 168
DB 7 ECANFIKYLVAANQCHLYACGTGAFHPTCTYLEVGHHEEDNIFKQDSHFEGRGSPYD 66
QY 169 AGHANVALFADGKYSATVDFPLADAVIYRSIGESPILRTYKHSKMLKEPFYQAV--- 225
DB 67 PFLTASLLIDELYSGTAAANMGDPALFRTLGHHPHPIRTQHSRWLNDRFISAHLI 126
QY 226 ----VDYGDYIYFFPREIAVEYNTWGVFPVAVQVCKNDMGSGQVLEKQWTSFLKARL 281

```

```

Db      127 PESDNPEDDKVYFFPRENAIDGESHGKATPHRIQICKNDDFGG-HRSIVNKKTTTLKARL 185
      282 NSGVG-----DSHFENILQAVTDVIRINGRD-----VLATSTSTYNSIPGSAVCAYMUL 333
      186 ICSVGPNGIDTHF-----DELQVFLMNSKDPKPIYGVFTTSSNIFKGSAAVCMYMS 240
Qy      334 DIASVFTGFKQKSPDSTWTVPDERVFKPRGCCAGSSSLERATSNPEPDITLFIK 393
      241 DVRVLILGVAHRDGNVQWVEY-QGRVYPRPGCP-SKTGSGDSTTDLDDVITFAR 298
Db      394 THPLMDEAVPSIENRPFRLTMVRYRLKIAVDTAAGYQNTVAVFLSGEKIILKFLAR 453
      299 SHPAMVNPVFINNRIIMIKTDVNOFTQIVDRAVDADGQYDAVEIGTDVETVLKVVSV 358
Qy      454 IONGSLNDSLFLEEMSYVNSKCSYDGEDRKIMQMDRASSSLVYAFSCVYKVPFG 513
      359 PRETHDLEEVLEENTVPR-----EPTTISAMELSTQOQYIGSTIGVAKQPLH 409
Db      514 RGRHCKCKKCTCIASRDYCGWIKEGACSHLSPNS-RLTFEODIERGNTDGLGDCHNSF 572
      410 RCDIYKACAECLADPYCAM--DSSCSRYPFTAKRTRRQDIRNG--DLTTCSDLQ 465
Qy      573 VALNGHSSSL---LPSTTSDSTAQEGYESRGMLDWKHLDDSPDSTPLGAVSHNQ 628
      466 HHNDHNGPSLEERIIYGVENSSTFLECSPEKSGALVYQFORNEDRKEEI-XMGDHILR 524
Db      629 DKKGVIRESYLK 640
Qy      525 TQGGTLRLSLQK 536

```

RESULT 11

hypothetical protein Y545B.1 - *Caenorhabditis elegans*
 C/Spectes: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_rev15ion 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T27165
 R:Jennard, N.
 submitted to the EMBL Data Library, October 1998
 A/Reference number: Z20321
 A/Accession: T27165
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-712 <NHL>
 A/Cross-references: EMBL:AL032653; PIDN:CAA21714.1; GSPDB:GN00019; CESP:Y545B.1
 A/Experimental source: clone Y545B
 C/Genetics:
 A/Gene: CESP:Y545B.1
 A/Map position: 1
 A/Intons: 36/1; 70/2; 85/3; 194/2; 221/1; 253/3; 301/1; 352/3; 522/2; 560/1; 599/3; 655

Query Match 14.5%; Score 789; DB 2; Length 712;
 Best Local Similarity 28.7%; Pred. No. 8.8e-45;
 Matches 229; Conservative 129; Mismatches 282; Indels 158; Gaps 29;

```

Db      6 LLLFTLLHFGAGFPDESEPISSHGNYTKOYF-----VFVGHKGRNTTORHRLDIQ 59
      7 LLLLFNVV-----RSSRAIT---GGVNLARKQILNSVIGIDRFGGIGTSDSDSDHF 55
Qy      60 MIMIMNG-TLYIARHDHITYVDIDTSHTEIYSGSKULTWRSROADVDTCKMKGKHDECH 118
      56 KLLAADDSILVGARNAVNLSTST-----LSVNHKIDMKPRPAHIEICIMKGSKTDCQ 110
Db      119 NFIYVLAKKDDALFVCGTNAFNPSCKNYKMDLLEPG---DEFGMARCPYDAKHANV 174
      111 NYIKVLARKSGVSLVCGTHAFSPKCRBY--TVTEFGIRNTRQFDQGSIPYPKINSS 167
Qy      175 ALFADG--KIYSATVTDFLAIDAVIYR-SLGESEPT-----LRTVKHDSKMLKEPYVOA 225
      168 ALIYVPGINQFLVAIVTDFVGVNDALIKRTIDETSSSAANIRFQSDARVLANFNFVAI 227
Db      226 VDYGDYIYFFPREIAVEY--NTMGKVVFPFRAVQYCKMDMGSGVLEKQMTSFLKALNC 283

```

```

Db      228 FAYKEHYVFWRELIASPAIDNNEPDIYARVAVCKXNDKGA-RPARNRTSYLKARLNC 286
      284 SVP-GDSHFYENILQAVTDVIRI-NGRDVLATSTSTYNSIPGSAVCAYMDIASVF-T 340
      287 SLPSGSSPFYENELKAVSDPIDAGNNHVVYTVSTPDSVDYRMAVACFKSMKKEEFPDN 346
Qy      341 GREFKQKSPDSTWTVPDERVFKPRGCCAGSSSLERATSNPEPDITLFIKHLMD 400
      347 GTFPHQNNAAQSMMAFNEFNEVPKPRPGSCSDST-----KLPENTVSFLHHPILHR 398
Db      401 AVPSIENRPFRLTMVRYRLKIAV---DTAAGPYQNTVAVFLSGEKIILKFLARIGNS 457
      399 PIPSV-AAPLLVEGADRADLTQITVLPRVAVGSH-NYDILFIGSDKVLKVEVDGNA 456
Qy      458 GFLND-SLFLSEMSYVNSKCSYDGEDRKIMQMDRASSSLVYAFSCVYKVPFGCE 516
      457 TVIQSATVFGQGVPLVN-----LTTKESVIVLSADEIASLPHNCA 498
Db      517 RHGKCKKCTCIASRDYCGWIKEGACSHLSPNSRLTFEODIERGNTDGLGDCHNSFV 576
      499 QOTSCK-CVQLODPHCAMDSSIARCVHGSWTGQFIQNVVFGSE--QCPBGITV-- 552
Qy      577 GHSSSLPSTTSDSTAQEGYESRGMLDWKHLDDSPDSTPLGAVSHNQDKGVIRE 636
      553 -----REVFDNSEAOPEAVS-----RS 571
Db      637 SYLKGHDLVPLYLLAIVLAIVMGAVPSGITVYCVCDHRKRDVAVVQRKKEILHS-- 694
      572 GYFKHSTITVLAIVASLISLIGA-FIGIRV-----NRWATSEPRSSASTSGSDY 625
Qy      695 -----RGSMSVTKLSLFGDQSKDPRKPAIILPLMHNKGKLTAPGTAKMLICA 745
      626 DSFGARALTRHDSLTTATKVDHF-----VPQSKQSVDAISLVMSIVA 668
Db      746 DQHDLITLTPRESTPT 763
Qy      665 THHPMSMGHSGSINTPS 686

```

RESULT 12

semaphorin II precursor - fruit fly (*Drosophila melanogaster*)
 C/Spectes: *Drosophila melanogaster*
 C/Date: 06-Jan-1995 #sequence_rev15ion 06-Jan-1995 #text_change 07-May-1999
 C/Accession: C49423
 R:Kolodkin, A.L.; Mathes, D.J.; Goodman, C.S.
 Cell 75, 1389-1399, 1993
 A/Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
 A/Reference number: A49423; MUID:94094312; PMID:8269517
 A/Accession: C49423
 A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A/Molecule type: mRNA
 A/Residues: 1-724 <KOL>
 A/Cross-references: GB:L26083
 C/Genetics:
 A/Gene: sema II
 A/Cross-references: FlyBase:FBgn0011260
 C/Superfamily: semaphorin

Query Match 14.3%; Score 779.5; DB 2; Length 724;
 Best Local Similarity 33.5%; Pred. No. 3.9e-44;
 Matches 193; Conservative 99; Mismatches 199; Indels 85; Gaps 22;

```

Qy      33 NYTKQYVVFYGHKGRNTTORHRLD-----IQMIMN---GLYIARHDHITY 79
      32 NFYYERPCCTGNDQGNMNYGKHGADVREPNCKLYRTFHMEDBDTLVGMADRVFRV 91
Db      80 DITSHTEIYSGK-KLTWRSROADVDTCKMKGKH-DECHNFIYLLKND-DALFVCG 136
      92 NLGNISSESN-CNRDAINLEPTDDVYSCVSKSGQLFDCKNHVRVIGSMDDQDRLVCG 149
Qy      137 TNAFNPSCKNY-----KMDLLEPGDEFSGMARCPYDAKHANVALLFADG----- 180

```

```

Db      150 TNAANP--KQVVIYANLTHLPRSEYVIGV---LGIAKCEYDPIDNSTALIVEANGNGL 204
Qy      181 -KLVSATVTPFLAIDAVIYRS-----IGSEPTLTIVKDSKWLKEPYFVQAVDYG 231
Db      205 PGLYSGTNAEFTKADTVIFRTDLYNTSAKRLEYKFKRTLKXDSKWLDPKPFVGSFDIG 264
Qy      232 IYFFFRFIAVYNNMGKVPFPRVAQVCKNDMGSKORULEKOWISFLKARLNCSPGSHF 291
Db      265 VYFFFRFIAVYNNMGKVPFPRVAQVCKNDMGSKORULEKOWISFLKARLNCSPGSHF 323
Qy      292 YENILOAVTVIRINGRDVILATPSTPYNS:PGSAVCAYMDLIDVAFVTSFGSKPKSPDS 351
Db      324 YFNEIQSYVQLPSKSR--PFAITFTSTNGLSGAVSGFHNEIQAAFNGKFKGQSSNS 381
Qy      352 TWTFVDERVYKPRPGCCAGSSSLERYATSNEFPDDTLNFKTHPLMDEAVPSIFNRP-W 410
Db      382 AMELVMSRVEPRPGCVNDTS-----NLDPVTILNFRSHPLMDKXVHNHPVY 433
Qy      411 FLRTMVRRLT--KIANDTAAGPIQNHVTVFLGSEKGIILKFLARIGNSGFLNLSLE 468
Db      434 YKRDVLPFKLVVDKIRIDIL--NQEYIVYVGTNLRIYKIVQYRNGESLSKLDIFE 490
Qy      469 MSVYNSEKSYDVEDKRIKMGMDLBRASSLSIYVAFSTCVIKVPIGRG-ERHGCCKTCTIA 527
Db      491 VA-----PNEAIQVMEISQTRKSLYIGTDHRIKOIDLAMCKRRINDCPR-CV- 536
Qy      528 SRDPYCGWIKEGGACSHLSPNSRLTFEODIERGNTD 563
Db      537 -RDPYCGMDKAMNCRPY----ELDLQDVANETS 567

```

RESULT 13

```

G02173
semaphorin III family homolog - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C/Accession: G02173
R.Naylor, S.
submitted to the EMBL Data Library, October 1995
A/Reference number: G09275
A/Accession: G02173
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-753 <N>
A/Cross-references: EMBL:U38276; NID:g1061350; PID:AA818276.1; PID:g1061351
C/Superfamily: semaphorin

```

```

Query Match      14.1%; Score 768; DB 2; Length 753;
Best local Similarity 31.6%; Pred. No. 2,4e-43;
Matches 191; Conservative 109; Mismatches 207; Indels 98; Gaps 22;

Qy      6 LLLYFTLLHFAAGPEDESEPSISHGNYTKQPVFVGHKPGKNTTQRRHLDIQMI----- 61
Db      6 LLLMSLLTGAMPSFTQD-----HLR---ATPRVRLSFKELKATG 43
Qy      62 -----MIMGST-----LYTAARDHIYTVIDTISHTEIYCSKLTAKSQA 102
Db      44 TAAHFNPLNTDYLILKKEDHDHMYGSKDYVSLDLHDIDNREPLI-----IHWAAPQ 99
Qy      103 DVDYTERMGKGI-KDSCHNFIKVLKKNDALFVCGTNAFNPSC-----RNYKMDL 152
Db      100 RLEECVLGKDVNGSCGFVRLIQPNRTHLYVCGTGAINEPCYVNRGRACQD-ITYL 158
Qy      153 EPPGDFSGMARCPYDAKXANALFADGKLVSATVDELAIDAVIYRSIGSPFLRTYKH 212
Db      159 EPERLE-SGKGCYDPKLDITASALINEELVAGVYIDFNGTDAALFRLTGKQTARTQY 217
Qy      213 DSKMLKEPYVQA-----VDYGVYIYFFFRFIAVYNNMGKVPFPRVAQVCKNDMGSK 266
Db      218 NSRMLNDPFIHAELIPDSAEKNDKLYFFFRSAAE-APQSPAVYARIGRIINDDGGHC 276
Qy      267 RYLEKOWISFLKARLNCSPGDS--HFYENILOAV---TVIRINGRDVILATPSTPYN 320

```

```

Db      277 CLVKN-MSFTLKRVLCSVPGDEGDIETHFDELQDVVQOQDVR---NPVIYAVFTSSGS 332
Qy      321 STPGSAVCAYMDLIDVAFVTSFGSKPKSPDSTTPVDEVEKPRPGCCAGSSSLERYAT 380
Db      333 VFRGSAVCYVSMADIMWFEPGPAHKEGPYQMPF-SGMPYPRPCTCGGFTTBSMK 391
Qy      381 SNEFPDNLNFKTHPLMDEAVPSIFNRPWFLRTMVRRLTKLA---VDTAAGPYQNHV 437
Db      392 TKQYDPEVIMFMSHPLMVAVPLQRRPLVATGAPYRLTTIAVDVDADGRYE---V 448
Qy      438 VFLGSEKGIILKFLARIGNSGFLNLSLFEEMSVYNSEKSYDVEDKRIKMGMDLBRASS 497
Db      449 LFIQTRGVYQYIV-LPXODQEMELMLEVEVEK-----DAPVKTMTISSKQ 498
Qy      498 SLIYASTCYIKVPLRCGERHGCCKTCTIASRDPYCGWIKEGGACSHLSPNS-RLTFEOD 556
Db      499 QLYVASAVGTLSLHRQCAVGAACDCLARDPYCAM--DGAQCSRYTASSKRRSRQD 556
Qy      557 IERGN 561
Db      557 VRGN 561

```

RESULT 14

```

148748
semaphorin E - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C/Accession: 148748
R.Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A/Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A/Reference number: 148744; MUID:95267431; PMID:7748561
A/Accession: 148748
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-751 <RES>
A/Cross-references: EMBL:X85994; NID:g854331; PID:CAA59986.1; PID:g854332
A/Genetics:
A/Gene: sem
C/Superfamily: semaphorin

```

```

Query Match      13.5%; Score 737; DB 2; Length 751;
Best local Similarity 31.3%; Pred. No. 2,9e-41;
Matches 187; Conservative 103; Mismatches 224; Indels 74; Gaps 18;

Qy      24 SEPSISHGNYTKQPVFVGHKPGKNTTQRRHLDIQMI--MNGTLYIARDHIYTVI 81
Db      42 SEYSLH-----QOLDYRILMDEODDRIYVGSKHILSLNI 79
Qy      82 DTSHTBELYCSKLTWKSROADVDTCKMKGHXDE--CHNFIKVLKKNDALFYCGTNAF 140
Db      80 NNISQEPV---SVFWPSTIKVECKWAGKDPHGCNFRVVIQTFNRTHLYVCGSGAF 135
Qy      141 NPSC-----RNYKMDLTFEFGDEF-----SGMARCEYDAKXANALFADGKLVSATV 188
Db      136 SPVCTLYNRGRSE-----DQVFMIDSKSGSGKRGCFNNVNTVSVMINELFSGWYI 189
Qy      189 DFLAIDAVIYRSIGSEPTLRTVHKDSKWLKEPYVQA-----VDYGV-YIYFFFRFIAV 241
Db      190 DFMGTDAALFRLSLFKRMQLRTDCHNSKWLSEPMFEDAVIBDGTDPNKAQVFFPKERLT 249
Qy      242 EYNMGKVPFPRVAQVCKNDMGSGQRYLEKOWISFLKARLNCSPGDS--HFYENILOAV 299
Db      250 DNNSTQKISHMILARICENDTGG--QRSLVNKMTITLKRVLCSYVDEDEGPETHFELDEV 308
Qy      300 TDVIRINGR-DVVLATPSTPYNS:PGSAVCAYMDLIDVAFVTSFGSKPKSPDSTVTPVD 358
Db      309 FLLETNDNRITLVVGIFTSSSVFSGSAVCYHSLDQIVENGPRPAHKEGPVHOLISY-Q 367
Qy      359 ERVPRPGCCAGSSSLERYATSNEFPDDTLNFKTHPLMDEAVPSIFNRPWFLRTMVR 418

```


Db 368 GRIPYRPGTCGGAFPTNNRTTKDFPDQVVTFRHNPMLNYSISPIHRRPLIVRIGTDY 427
 QY 419 RLTKIADVAAGPYQHNTVFLGSEKGIILKFLARIGNSGFLNDSLFREMSVYNSKES 478
 Db 428 KYTKIADRVNADGKYNHFLGTDRGTQVKVVLPTNSS-ASGELLLELEVFKNH--- 483
 QY 479 YDGEDKRMGMQLDBASSSLYYAFSTCVIKVPLGRCERHKKCKTKCIASRDPYCGWIKI 538
 Db 484 -----VPIITMEISSKKQQLYVSNBGSQVSLHCHITGTACADCCCLARDPYCAM--D 535
 QY 539 GCACSHLP-NSRLTFEODIERGNTDGLDCHNSFVALNGHSSSLPSTTSDSTQAE 595
 Db 536 GHCSRFYPTGKRSSRDVVRHGNP--LTQCRGFNLKAYRNAETIVQYGVRRNSTPLE 591
 RESULT 15
 S66498
 M-sema F protein precursor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #ext_change 24-Nov-1999
 C/Accession: S66498
 R/Inagaki, S.; Furuyama, T.; Iwahashi, Y.
 FEBS Lett 370, 269-272, 1995
 A/Title: Identification of a member of mouse semaphorin family.
 A/Reference number: S66498; PMID:95385809; PMID:7656991
 A/Accession: S66498
 A/Status: Preliminary
 A/Molecule type: mRNA
 A/Residues: 1-834 <TNA>
 A/Cross-references: EMBL:S79463; NID:g1110598; FIDN:AA35184.1; PID:g1110599
 C/Superfamily: semaphorin
 F/1-21/Domain: signal sequence #status predicted <SIG>
 F/22-834/Product: M-sema F protein #status predicted <MAT>
 Query Match 12.7%; Score 692; DB 2; Length 834;
 Best Local Similarity 26.0%; Pred. No. 3,5e-36;
 Matches 232; Conservative 131; Mismatches 326; Indels 204; Gaps 34;
 QY 65 NGTLIYARDHIYVIDIDTSHTEIYSCKLITWKSROADVDTCRMKGK-HKDECHNFIKY 123
 Db 62 SGLLYVAREHLEAFSV-----EALHLOGAISWEAPAEKKECTQKGSNQTECFNFI 116
 QY 124 LLLKNDALFYCGTNAENPSCRYKMDTLBPFGDEF-SGMARCPYDAKHNVALFADGKL 182
 Db 117 LQPNSSHLYVCGTYAFQPKCTYINMLFTLDRAEFEDGKCPYDPAGKHTGLVDGEL 176
 QY 183 YSATVTEFLAIDAVYSLGSEPTLRVTKHDSKWLKEPYVQA-----VDYGDYI 232
 Db 177 YSATLNNFLGTBPIILMYGTHHSIKT-EYIAFWLNEPHFVGSAPFVESVSGFTGDDXI 235
 QY 233 YFFPREIAVEYNTMGKVPFRAVQYCKNDWGSQVLEKOWTSFLKARLNCVPGDSHFY 292
 Db 236 YFFPREIAVEYDCYSEGVAVARVCKDMGGA-RTLOKKWTFPLKARLVCAP-DWKVY 293
 QY 293 FNIIQATVDVIRINGRVV-LATSTPYNSIPGSANCAVMDLDIASVFTGRFKEQKSPDS 351
 Db 294 FNQKAVHTLFGASWHHTTFEGVFQARWGMDDLSAVCEYQLEQVVEGPKYKESQAQ 353
 QY 352 TWTVDPDERVPRPGCGAGSSSLER-YATSNPEPDDTLNFIKTHPLMDEAVPSIFNRPW 410
 Db 354 KMAKRTDB-VSPSPGSGINWHRDNGTSSLELPDNTLNFIKTHPLMEDQVPRRLGRPL 412
 QY 411 FLRTVRYRLTKIADVTAAG-PYQNTVPLVSGSEKGIILKFLARIGNSGFLNDSL----- 464
 Db 413 LVKKNTF--THVADRVPLDGAITYVLFIGTGMILKAV-----SLGPMIH 459
 QY 465 FLEMSVYNSKESYDGEDKRMGMQLDBASSSLYYAFSTCVIKVPLGRCERHKKCKTKI 524
 Db 460 MVEELQVFDDEP-----VESLVSGSKVPLFAGSRQSLVQLSLADCTKYRFC-VD 508
 QY 525 CIASRDPYCGWIKESGAC-----SHLS--FNSRLTFEODIER----- 559
 Db 509 CVIARDPYCANVNTSRCVATTSGRSGSEFLQVHANLDTSGKCNQYGIKKVRSIPKNTIV 568

QY 560 -GNTDGLDCHNSFVALNGH---SSSLPSTTSDSTQAEYSGMGLDMKHLNDSPTS 615
 Db 569 VSGTDLVLPCHLSSNLAAHMTFGSQDLPAEQPGSFLYDTGLQALVW----- 616
 QY 616 TDPGLAVASHN-----HDKKG--VIRESYLK-----HDQVPTLLAIAVILAFV 660
 Db 617 ---AAQRHSGSPRYCYSEEGQTRLAASYLVAVAGSVTLBARAPLENLGLVWLAVVA 672
 QY 661 MGAVFSGITVYCVCDHRRKDYAVVQKKEKELTHSRGSSSVYKLSGLFGDTQSKDPKE 720
 Db 673 LGAVCL-VULLLVLLRRR-----LRELE-----KGAAS-----ER 704
 QY 721 ALLTPMENGKLATPGNTAKMLIKADQHHLDTALPTPESTTTLQCKRPSRSGRPM 777
 Db 705 TLVYPL-----ELPKEDASPPF--RPGPETDEKLMDEPV 735
 QY 778 -----ERNONLINACTKQMPWMSGPVIPNDLPLRASPSHIFSVVLPITQCGYQHEYVDQ 832
 Db 736 GYYSDGSLKIVPGHARCPGSGGSPSPPGITGQPLPS--PTRLHGGGRNSNMANQYR- 792
 QY 833 PKMSEVAQVALDQAAATLEKTIKEHLSSKSPNHGVNLENLDSLPPKVPQRE 885
 Db 793 -----LQLGSDRGS-----GHPLBELADELRKXQCRQ 822

Search completed: October 23, 2003, 17:11:31
 Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 23, 2003, 17:10:42 ; Search time 84 Seconds
(without alignments)
2053.386 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450
Sequence: 1 MRSALLLYFTLLHFAAGCF.....PPKPSFAPLSTSMKNDACT 1030

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5422.5	99.5	1047	11	US-09-957-187-85
2	4931	90.3	939	11	US-09-957-187-4
3	4590.5	84.2	884	11	US-09-957-187-6
4	3373	61.9	630	11	US-09-957-187-30
5	3362	61.7	626	11	US-09-957-187-83
6	2013.5	36.9	888	12	US-10-391-413-4
7	2009.5	36.9	888	12	US-09-931-836-35
8	2009.5	36.9	888	12	US-10-035-977-35
9	2009.5	36.9	888	12	US-10-137-870-544
10	2009.5	36.9	888	12	US-10-140-018-544
11	2009.5	36.9	888	12	US-10-140-021-544
12	2009.5	36.9	888	12	US-10-140-274-544
13	2009.5	36.9	888	12	US-10-140-471-544
14	2009.5	36.9	888	12	US-10-140-807-544
15	2009.5	36.9	888	12	US-10-140-922-544

15	2009.5	36.9	888	12	US-10-140-922-544	Sequence 544, App
17	2009.5	36.9	888	12	US-10-140-926-544	Sequence 544, App
18	2009.5	36.9	888	12	US-10-141-698-544	Sequence 544, App
19	2009.5	36.9	888	12	US-10-141-702-544	Sequence 544, App
20	2009.5	36.9	888	12	US-10-141-704-544	Sequence 544, App
21	2009.5	36.9	888	12	US-10-142-421-544	Sequence 544, App
22	2009.5	36.9	888	12	US-10-142-432-544	Sequence 544, App
23	2009.5	36.9	888	12	US-10-142-767-544	Sequence 544, App
24	2009.5	36.9	888	12	US-10-143-033-544	Sequence 544, App
25	2009.5	36.9	888	12	US-10-144-994-544	Sequence 544, App
26	2009.5	36.9	888	12	US-10-145-628-544	Sequence 544, App
27	2009.5	36.9	888	12	US-10-145-631-544	Sequence 544, App
28	2009.5	36.9	888	12	US-10-145-743-544	Sequence 544, App
29	2009.5	36.9	888	12	US-10-145-746-544	Sequence 544, App
30	2009.5	36.9	888	12	US-10-145-748-544	Sequence 544, App
31	2009.5	36.9	888	12	US-10-145-823-544	Sequence 544, App
32	2009.5	36.9	888	12	US-10-145-826-544	Sequence 544, App
33	2009.5	36.9	888	12	US-10-145-870-544	Sequence 544, App
34	2009.5	36.9	888	12	US-10-145-876-544	Sequence 544, App
35	2009.5	36.9	888	12	US-10-145-895-544	Sequence 544, App
36	2009.5	36.9	888	12	US-10-146-724-544	Sequence 544, App
37	2009.5	36.9	888	12	US-10-146-725-544	Sequence 544, App
38	2009.5	36.9	888	12	US-10-146-795-544	Sequence 544, App
39	2009.5	36.9	888	12	US-10-147-495-544	Sequence 544, App
40	2009.5	36.9	888	12	US-10-147-501-544	Sequence 544, App
41	2009.5	36.9	888	12	US-10-147-504-544	Sequence 544, App
42	2009.5	36.9	888	12	US-10-147-506-544	Sequence 544, App
43	2009.5	36.9	888	12	US-10-147-509-544	Sequence 544, App
44	2009.5	36.9	888	12	US-10-147-510-544	Sequence 544, App
45	2009.5	36.9	888	12	US-10-147-511-544	Sequence 544, App

ALIGNMENTS

RESULT 1
US-09-957-187-85
Sequence 85, Application US/09957187
Publication NO. US2003054514A1
GENERAL INFORMATION:
APPLICANT: Shimkets, Richard A.
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 15966-540 CIP
CURRENT FILING DATE: 2000-09-19
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 60/123,667
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/234,082
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,798
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/174,485
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 85
LENGTH: 1047
TYPE: PRT
ORGANISM: Homo sapiens
US-09-957-187-85
Query Match 99.5%; Score 5422.5; DB 11; Length 1047;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1029; Conservative 0; Mismatches 1; Indels 17; Gaps 1;
QY 1 MRSALLLYFTLLHFAAGCFPEDESPISISHGNYTKQYVFVGHKPGRRNTTORHRLDIOM 60
DB 1 MRSALLLYFTLLHFAAGCFPEDESPISISHGNYTKQYVFVGHKPGRRNTTORHRLDIOM 60
QY 61 IMINNGTYIARADHIYTVDDTSHTEETIYCSKLTWKSROADVDTCRMKGHKXDCNHF 120

Db 61 IMIMNTLYIAARDHIYTDIDTSHTEETIYCSKLTWKSROADVDCRMKGKHCCHNF 120
 QY 121 IKVLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDESSGMARCPDYAKHANVALFADG 180
 Db 121 IKVLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDESSGMARCPDYAKHANVALFADG 180
 QY 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKMLKEPPYQAVDYGDIYFFPREIA 240
 Db 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKMLKEPPYQAVDYGDIYFFPREIA 240
 QY 241 VEYNTMGKVPFPPVAQVCKNDMGSGQRYLEKQMTSFLKARLNCVGDSHFYFNILQAVT 300
 Db 241 VEYNTMGKVPFPPVAQVCKNDMGSGQRYLEKQMTSFLKARLNCVGDSHFYFNILQAVT 300
 QY 301 DVIRINGRDVLTATFSTPYNISIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 Db 301 DVIRINGRDVLTATFSTPYNISIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 QY 361 VPKRPPCCAGSSSLERYATSNFPDPTLNFITHPMDAVSIFNRPWFLLTMVRYRL 420
 Db 361 VPKRPPCCAGSSSLERYATSNFPDPTLNFITHPMDAVSIFNRPWFLLTMVRYRL 420
 QY 421 TKTAVDTAAPRYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFILEMSVYNSEKSYD 480
 Db 421 TKTAVDTAAPRYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFILEMSVYNSEKSYD 480
 QY 481 GVEDKRMQDLBRASSSLVYAFSTCVIKVPLGRCERHCKCKTCLASRDPYCGMIKEGG 540
 Db 481 GVEDKRMQDLBRASSSLVYAFSTCVIKVPLGRCERHCKCKTCLASRDPYCGMIKEGG 540
 QY 541 ACSHLSNRSRLTFEODIERGNTDGLGCHNSFVALN-----GHSSSL 583
 Db 541 ACSHLSNRSRLTFEODIERGNTDGLGCHNSFVALN-----GHSSSL 583
 QY 584 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 643
 Db 584 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 643
 QY 601 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 660
 Db 601 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 660
 QY 644 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 703
 Db 644 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 703
 QY 661 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 720
 Db 661 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 720
 QY 704 KLSGLFSDTOSKDPKPAIILTPLMHNGKLATPGNTAKMLIKADQHHLLDTRLPPESTPT 763
 Db 704 KLSGLFSDTOSKDPKPAIILTPLMHNGKLATPGNTAKMLIKADQHHLLDTRLPPESTPT 763
 QY 721 KLSGLFSDTOSKDPKPAIILTPLMHNGKLATPGNTAKMLIKADQHHLLDTRLPPESTPT 780
 Db 721 KLSGLFSDTOSKDPKPAIILTPLMHNGKLATPGNTAKMLIKADQHHLLDTRLPPESTPT 780
 QY 764 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVLLPTTQ 823
 Db 764 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVLLPTTQ 823
 QY 781 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVLLPTTQ 840
 Db 781 LQCKRKRSRGRERERONLINACTKMPMGSPVITPDLPLRASPSHISBVVLLPTTQ 840
 QY 824 GYQHEVYDQPKMSVAVQWALEDDQATLETKIKELHLSKSPHNGVNLVENLDSLPKVPQ 883
 Db 824 GYQHEVYDQPKMSVAVQWALEDDQATLETKIKELHLSKSPHNGVNLVENLDSLPKVPQ 883
 QY 841 GYQHEVYDQPKMSVAVQWALEDDQATLETKIKELHLSKSPHNGVNLVENLDSLPKVPQ 900
 Db 841 GYQHEVYDQPKMSVAVQWALEDDQATLETKIKELHLSKSPHNGVNLVENLDSLPKVPQ 900
 QY 884 REASLGPBGASLSQTCISKLLEHSHSSSYVDYKSYPTNSLTSHQATTLKRNNTSSN 943
 Db 884 REASLGPBGASLSQTCISKLLEHSHSSSYVDYKSYPTNSLTSHQATTLKRNNTSSN 943
 QY 901 REASLGPBGASLSQTCISKLLEHSHSSSYVDYKSYPTNSLTSHQATTLKRNNTSSN 960
 Db 901 REASLGPBGASLSQTCISKLLEHSHSSSYVDYKSYPTNSLTSHQATTLKRNNTSSN 960
 QY 944 SSHLSRNGSPGRGNPPAPQPVDSIOVHSSQPSGQAVTVSRQPSLNNYNSLTSGKRT 1003
 Db 944 SSHLSRNGSPGRGNPPAPQPVDSIOVHSSQPSGQAVTVSRQPSLNNYNSLTSGKRT 1003
 QY 961 SSHLSRNGSPGRGNPPAPQPVDSIOVHSSQPSGQAVTVSRQPSLNNYNSLTSGKRT 1020
 Db 961 SSHLSRNGSPGRGNPPAPQPVDSIOVHSSQPSGQAVTVSRQPSLNNYNSLTSGKRT 1020
 QY 1004 PSIKPDVPPKPSFAPLSTSMKPNDACT 1030
 Db 1004 PSIKPDVPPKPSFAPLSTSMKPNDACT 1030
 QY 1021 PSIKPDVPPKPSFAPLSTSMKPNDACT 1047
 Db 1021 PSIKPDVPPKPSFAPLSTSMKPNDACT 1047

; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: laRoche, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-540 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/174,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 4
 ; LENGTH: 939
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-957-187-4
 Query Match 90.3%; Score 4921; DB 11; Length 939;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 928; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MRESEALLTYFLIHFAAGFPEDSEPISTSHGNYTQYVFPVGHKRGNTTQNRHLDIOM 60
 Db 1 MRESEALLTYFLIHFAAGFPEDSEPISTSHGNYTQYVFPVGHKRGNTTQNRHLDIOM 60
 QY 61 IMIMNTLYIAARDHIYTDIDTSHTEETIYCSKLTWKSROADVDCRMKGKHCCHNF 120
 Db 61 IMIMNTLYIAARDHIYTDIDTSHTEETIYCSKLTWKSROADVDCRMKGKHCCHNF 120
 QY 121 IKVLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDESSGMARCPDYAKHANVALFADG 180
 Db 121 IKVLKKNDDALFVCGTNAFNPSCRYKMDTLEPFGEDESSGMARCPDYAKHANVALFADG 180
 QY 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKMLKEPPYQAVDYGDIYFFPREIA 240
 Db 181 KLYSATVTDPLAIDAVIYISLGSPTLRKYKHSKMLKEPPYQAVDYGDIYFFPREIA 240
 QY 241 VEYNTMGKVPFPPVAQVCKNDMGSGQRYLEKQMTSFLKARLNCVGDSHFYFNILQAVT 300
 Db 241 VEYNTMGKVPFPPVAQVCKNDMGSGQRYLEKQMTSFLKARLNCVGDSHFYFNILQAVT 300
 QY 301 DVIRINGRDVLTATFSTPYNISIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 Db 301 DVIRINGRDVLTATFSTPYNISIPGSAVCAVMDLDIASVFTGRFKEQKSPDSTWTPVDER 360
 QY 361 VPKRPPCCAGSSSLERYATSNFPDPTLNFITHPMDAVSIFNRPWFLLTMVRYRL 420
 Db 361 VPKRPPCCAGSSSLERYATSNFPDPTLNFITHPMDAVSIFNRPWFLLTMVRYRL 420
 QY 421 TKTAVDTAAPRYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFILEMSVYNSEKSYD 480
 Db 421 TKTAVDTAAPRYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFILEMSVYNSEKSYD 480
 QY 481 GVEDKRMQDLBRASSSLVYAFSTCVIKVPLGRCERHCKCKTCLASRDPYCGMIKEGG 540
 Db 481 GVEDKRMQDLBRASSSLVYAFSTCVIKVPLGRCERHCKCKTCLASRDPYCGMIKEGG 540
 QY 541 ACSHLSNRSRLTFEODIERGNTDGLGCHNSFVALN-----GHSSSL 583
 Db 541 ACSHLSNRSRLTFEODIERGNTDGLGCHNSFVALN-----GHSSSL 583
 QY 584 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 643
 Db 584 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 643
 QY 601 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 660
 Db 601 PSTTSTSTOEGESRGMMDMKHLLDSPDTPGLAVSSHNDKKGVIRESYLYKGD 660
 QY 644 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 703
 Db 644 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 703
 QY 661 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 720
 Db 661 QLVAVTLLAIAVILAFMGAVFSGITVYCVCDHRRKDVAVVORKEKELTHSRGSMGSVT 720

Db 661 MGAFSGITTCVCDHRRKDAVAVQKKEKELTHSRGMSVTKLSGLFGDTQSKDPKE 720
 Qy 721 AITPLMHNGKLATPGNTAKMLIKADQHHLDTLALPTPESTPTLQOKRPSRGSEMERN 780
 Db 721 AITPLMHNGKLATPGNTAKMLIKADQHHLDTLALPTPESTPTLQOKRPSRGSEMERN 780
 Qy 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVYVLPITQOGYQHEYVQPKMSEVAQ 840
 Db 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVYVLPITQOGYQHEYVQPKMSEVAQ 840
 Qy 841 MALEDOATLEYKTIKEHLSKSPNHGYNLVENLDSLPKYPQREASLGPPGASLSQTGL 900
 Db 841 MALEDOATLEYKTIKEHLSKSPNHGYNLVENLDSLPKYPQREASLGPPGASLSQTGL 900
 Qy 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHQAT 932
 Db 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHLT 932

RESULT 3
 US-09-957-187-6
 ; Sequence 6, Application US/09957187
 ; Publication No. US20030054514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shinketsu, Richard A.
 ; APPLICANT: Lachocelle, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-540 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/174,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 884
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-957-187-6

Query Match 84.2%; Score 4590.5; DB 11; Length 884;
 Best Local Similarity 93.7%; Pred. No. 0;
 Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;
 Qy 1 MRSSEALLTYFTLTHFAGFPEDESEFISISHQNTYKQYVFEVGHGREGNTTQRHRLDQM 60
 Db 1 MRSSEALLTYFTLTHFAGFPEDESEFISISHQNTYKQYVFEVGHGREGNTTQRHRLDQM 60
 Qy 61 IMINGGLYIARDHIYTVIDTSHTEELYCSKLTWKSROADVDTCRMKGKXDECFNF 120
 Db 61 IMINGGLYIARDHIYTVIDTSHTEELYCSKLTWKSROADVDTCRMKGKXDECFNF 120
 Qy 121 IYLLKNDALFVCGTNAFNSCENYKMDTLEPPDEDESSGARCGYDAKXANVALLFADG 180
 Db 121 IYLLKNDALFVCGTNAFNSCENYKMDTLEPPDEDESSGARCGYDAKXANVALLFADG 180
 Qy 181 KLYSATVDFLADAVYASLSGSPFLRTVKGDSKMLKEPYVOADVGYDYIYFFERELA 240
 Db 181 KLYSATVDFLADAVYASLSGSPFLRTVKGDSKMLKEPYVOADVGYDYIYFFERELA 240
 Qy 241 VEYNTMGKVVFPRAVQVCKNDMGSGQVLEKQWTSFKARLNCVPGDSHFFYFNILQAVT 300
 Db 241 VEYNTMGKVVFPRAVQVCKNDMGSGQVLEKQWTSFKARLNCVPGDSHFFYFNILQAVT 300

Qy 301 DVIRINGRDVVLATFTSPYNSIPGSACAYDMLDIASVFGFPECKSPDSTWTVPEDER 360
 Db 301 DVIRINGRDVVLATFTSPYNSIPGSACAYDMLDIASVFGFPECKSPDSTWTVPEDER 360
 Qy 361 VPKRPPCCAGSSGLERYATSNEDPDTLPIKTHPLMDEAVPSIENRPMFLRMVRYL 420
 Db 361 VPKRPPCCAGSSGLERYATSNEDPDTLPIKTHPLMDEAVPSIENRPMFLRMVRYL 420
 Qy 421 TKIADVTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFINDSLFLEMSVYNSEKCSYD 480
 Db 421 TKIADVTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGFINDSLFLEMSVYNSEKCSYD 480
 Qy 481 GVEDKRMGMQLDRASSSLVAFSTCVIKYPLGCEHSGKCKTCTIASRDPYCCMIKRG 540
 Db 481 GVEDKRMGMQLDRASSSLVAFSTCVIKYPLGCEHSGKCKTCTIASRDPYCCMIKRG 540
 Qy 541 ACSHLSNSRLTFEQDIERGNTDGLGDCNHSFVALNGHSSSLPSTTTSSTAOGEYSR 600
 Db 541 ACSHLSNSRLTFEQDIERGNTDGLGDCNHSFVALN----- 576
 Qy 601 GGMIDWKHLDSPESTPLGAVSHNODKKGVRESYIKGHDOLVPTLLAIVTLAFV 660
 Db 601 GGMIDWKHLDSPESTPLGAVSHNODKKGVRESYIKGHDOLVPTLLAIVTLAFV 660
 Qy 661 MGAFSGITTCVCDHRRKDAVAVQKKEKELTHSRGMSVTKLSGLFGDTQSKDPKE 720
 Db 661 MGAFSGITTCVCDHRRKDAVAVQKKEKELTHSRGMSVTKLSGLFGDTQSKDPKE 720
 Qy 721 AITPLMHNGKLATPGNTAKMLIKADQHHLDTLALPTPESTPTLQOKRPSRGSEMERN 780
 Db 721 AITPLMHNGKLATPGNTAKMLIKADQHHLDTLALPTPESTPTLQOKRPSRGSEMERN 780
 Qy 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVYVLPITQOGYQHEYVQPKMSEVAQ 840
 Db 781 QNLINACTKMPMPGSPVITPDLPLRASPSHPSVYVLPITQOGYQHEYVQPKMSEVAQ 840
 Qy 841 MALEDOATLEYKTIKEHLSKSPNHGYNLVENLDSLPKYPQREASLGPPGASLSQTGL 900
 Db 841 MALEDOATLEYKTIKEHLSKSPNHGYNLVENLDSLPKYPQREASLGPPGASLSQTGL 900
 Qy 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHQAT 932
 Db 901 SKRLEMHSSSYGVDYKRSYPTNSLTRSHLT 877

RESULT 4
 US-09-957-187-30
 ; Sequence 30, Application US/09957187
 ; Publication No. US20030054514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shinketsu, Richard A.
 ; APPLICANT: Lachocelle, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-540 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/174,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 30
 ; LENGTH: 630
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-957-187-30

Query Match 61.9%; Score 3373; DB 11; Length 630;
 Best Local Similarity 99.8%; Pred. No. 6.3e-275;
 Matches 628; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 17 GAGFPDSEPISSHGNYTKQYVFGHKGKGRNTTORHLDIQIMIMNGTLYIAARDHI 76
   1 GSGFPDSEPISSHGNYTKQYVFGHKGKGRNTTORHLDIQIMIMNGTLYIAARDHI 60
Db 77 YVVDITSHTEIYCSKLTWKSQRQADVDTCKMKGKHDECHNFIVLLKKNDALFVCG 136
   61 YVVDITSHTEIYCSKLTWKSQRQADVDTCKMKGKHDECHNFIVLLKKNDALFVCG 120
QY 137 TAAFNSSCRNYKMDLTPFGDESGMARCPYDAKHANVALFADGKLXSAVTDFLAIDAV 196
   121 TAAFNSSCRNYKMDLTPFGDESGMARCPYDAKHANVALFADGKLXSAVTDFLAIDAV 180
Db 197 IYRSLGESPTLRVKHDSKWLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQ 256
   181 IYRSLGESPTLRVKHDSKWLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQ 240
QY 257 VCKNDGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 316
   241 VCKNDGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 300
Db 317 TPNYSTPGSAVCAVMDLIDIASVFTGRFKEQKSPDSTWTPVDERVPRKPRGCCAGSSLE 376
   301 TPNYSTPGSAVCAVMDLIDIASVFTGRFKEQKSPDSTWTPVDERVPRKPRGCCAGSSLE 360
QY 377 RYATSNFPDDTLNFKTHPLMDAVALPSIFNRPMFLIRYRLTKIAVDIAAGPYQNH 436
   361 RYATSNFPDDTLNFKTHPLMDAVALPSIFNRPMFLIRYRLTKIAVDIAAGPYQNH 420
Db 437 VVFLESEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMQDLRAS 496
   421 VVFLESEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMQDLRAS 480
QY 497 SSIYVAFSTCVIVKPLGRCEHKGCKKCTCIASRDPCGMKEGACSHLSPNSRLTFEED 556
   481 SSIYVAFSTCVIVKPLGRCEHKGCKKCTCIASRDPCGMKEGACSHLSPNSRLTFEED 540
Db 557 IERGNTDGLGDCNHSFVALNGHSSSLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 616
   541 IERGNTDGLGDCNHSFVALNGHSSSLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 600
QY 617 DPLGAVSSHHQDKKGVIRESYLKGHDQ 645
   601 DPLGAVSSHHQDKKGVIRESYLKGHDQ 629
Db

```

RESULT 5
 US-09-957-187-83
 ; Sequence 83, Application US/09957187
 ; Publication No. US200300545141
 ; GENERAL INFORMATION:
 ; APPLICANT: Shimketa, Richard A.
 ; APPLICANT: Larochelle, William
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
 ; FILE REFERENCE: 15966-340 CIP
 ; CURRENT APPLICATION NUMBER: US/09/957,187
 ; CURRENT FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/123,667
 ; PRIOR FILING DATE: 1999-03-09
 ; PRIOR APPLICATION NUMBER: 09/520,781
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/234,082
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: 60/233,798
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 60/114,485
 ; PRIOR FILING DATE: 2000-01-04
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 83
 ; LENGTH: 626
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-957-187-83

Query Match 61.7%; Score 3362; DB 11; Length 626;
 Best Local Similarity 100.0%; Pred. No. 5.2e-274;
 Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 19 GEPDESEPISSHGNYTKQYVFGHKGKGRNTTORHLDIQIMIMNGTLYIAARDHI 78
   1 GEPDESEPISSHGNYTKQYVFGHKGKGRNTTORHLDIQIMIMNGTLYIAARDHI 60
Db 79 VVDITSHTEIYCSKLTWKSQRQADVDTCKMKGKHDECHNFIVLLKKNDALFVCGTN 138
   61 VVDITSHTEIYCSKLTWKSQRQADVDTCKMKGKHDECHNFIVLLKKNDALFVCGTN 120
QY 139 AFNPSCRNYKMDLTPFGDESGMARCPYDAKHANVALFADGKLXSAVTDFLAIDAV 198
   121 AFNPSCRNYKMDLTPFGDESGMARCPYDAKHANVALFADGKLXSAVTDFLAIDAV 180
Db 199 RSLGESPTLRVKHDSKWLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQVC 258
   181 RSLGESPTLRVKHDSKWLKEPFYVQAVDYGDIYFFPREIAVEYNTMGKVFPRVAQVC 240
QY 259 KNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 318
   241 KNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVTDVIRINGRDVLAITS 300
Db 319 YNSIFGSAVCAVMDLIDIASVFTGRFKEQKSPDSTWTPVDERVPRKPRGCCAGSSLE 378
   301 YNSIFGSAVCAVMDLIDIASVFTGRFKEQKSPDSTWTPVDERVPRKPRGCCAGSSLE 360
QY 379 AFSNFPDDTLNFKTHPLMDAVALPSIFNRPMFLIRYRLTKIAVDIAAGPYQNH 438
   361 AFSNFPDDTLNFKTHPLMDAVALPSIFNRPMFLIRYRLTKIAVDIAAGPYQNH 420
Db 439 FLGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMQDLRAS 498
   421 FLGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYDGEDKRIKMQDLRAS 480
QY 499 IYVAFSTCVIVKPLGRCEHKGCKKCTCIASRDPCGMKEGACSHLSPNSRLTFEED 558
   481 IYVAFSTCVIVKPLGRCEHKGCKKCTCIASRDPCGMKEGACSHLSPNSRLTFEED 540
Db 559 RGTDTGLGDCNHSFVALNGHSSSLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 618
   541 RGTDTGLGDCNHSFVALNGHSSSLPSTTSDSTAQEGYESRGMLDMKHLDSPDST 600
QY 619 LGAVSSHHQDKKGVIRESYLKGHDQ 644
   601 LGAVSSHHQDKKGVIRESYLKGHDQ 626
Db

```

RESULT 6
 US-10-391-413-4
 ; Sequence 4, Application US/10391413
 ; Publication No. US20030167482A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMURA, Toru et al.
 ; TITLE OF INVENTION: NOVEL SEKAPHORIN Z AND GENE ENCODING THE SAME
 ; FILE REFERENCE: 0020-5120P
 ; CURRENT APPLICATION NUMBER: US/10/391,413
 ; CURRENT FILING DATE: 2003-03-19
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 888
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-391-413-4

Query Match 36.9%; Score 2013.5; DB 12; Length 888;
 Best Local Similarity 45.3%; Pred. No. 3.4e-160;
 Matches 424; Conservative 129; Mismatches 241; Indels 141; Gaps 20;

5 ALLIYFLHLHAGAGPEDESEPIISHGNTYKQYVFEVGEKPGNTTOR--HRIDIMIM 62
 12 ALLIILLGAGHGLPEDEPPLISVAPRDYLNHYFVVGSGPRLTPEAGADNDIQRVL 71
 63 IMNTIYIARNDHIYTDIDTSHREIYCSKGLTWKSRQDVDCRKNKMGKCHNFIF 122
 72 RVNRTLTIGBDNLYRVELEPTSTELRYQKRLWRNSPDINVCNMGKQEGECNPFVK 131
 123 VLLKNDALFVCGTNAFNSCRNVMKDTLEPFDESGMARCPCYDAKANVALFADGKL 182
 132 VLLRDEBTLFVCGSNMFPNCANYSIDTLPQVDNDSGMARCYDDKHNVALFSGML 191
 183 YSATVTEFLAIDAVIYRSLGSSPTLRIVKHD SKMLKEPYQAVDYGDIYFFREIAYE 242
 192 FTATVTEFLAIDAVIYRSLGDRPTLRIVKHD SKMKEPYQAVDYGDIYFFREIAYE 251
 243 YNMGKVFYFVAVQVCKRDKGSGQVLEKQMTSFLKRLNCSVPGDSHFYFNILQAVTDV 302
 252 FNYLEKVVSVARCKNDVGSSPRVLEKQMTSFLKRLNCSVPGDSHFYFNILQAVTV 311
 303 IIRGRADVLTFTSPYNSIGSAVCAVMDIASVFTGRPEKQSPDWTVPYDERVP 362
 312 VSLGRFVVLAVFTSPNSIGSAVCAFDLQVNAVAFGRERQKSPSINTVPEDQVP 371
 363 KPRGCCAGSSSLERATSNFPPDTLNFIKTHPLMDEAVPSINRPFILTMVRYRLTK 422
 372 RPRGCCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSILGAPWILRTLRHQLTR 429
 423 IAVDTAAPRYONHYVFLGSEKGIILKELAR--IGNSGFLNDSLPLEMSVYNGEKSYD 480
 430 VAVDGAAPWGNQYTVFLSAGVTLKFLVPRNASTSGTSLVLEFEYTRPRCRGP 489
 481 GVED--KRMGMQDPRASSLYVAFSTCVIKVPLRCRCHKCKCTIASHRDPYCGWIK 538
 480 GGERGQGLISLELDASGGLIAPRCVAVPVAVRCQYSGCMNCIGSGDPYCGWAPD 549
 539 GGACSHLSNSRLTFEODIERGNTDGLDCHNSFVALNGHSSSLIPSTTSDSTAQEGYE 598
 550 -GSCIFLSPGTRAAPEQDVSGASTGLGDC----- 578
 559 SRGMLDMKHLIDS DSTDPGLGAVSHNHQCKKQYIRESYVKGHQLVPTLLATAVILA 658
 579 -----TGLRLASLEDRAGLVSVMILVTSSVAA 606
 659 FVMGAVFSGITV-YGVCDHRRKDVAVVORKEKE--LTSRSGSMSSVTKLSGLFQDTQK 715
 607 FVVGAVVSGFSGVGVGLRERELA--RRKDEALILAGAGAVLSVSRLL--GERAAQ 660
 716 DP-----KPEALITPLMHNGKLAIPGNTAK--MLIKADQHLIDLTALPTPEST 761
 661 GPGGRRGGGGGAGVPPREALIAPLMONG-----WAKATLLQGGPHIDISGLTPPEQT 713
 762 PTLQCKRR-----SGSRREMNQNLINACTKDMFWSPVITDLP--LAPSSH 811
 714 P-LPQKRLPTPHPHALGPRAMDH-----GPHLLPASASSLLLLAPAR 757
 812 IPSVVVLTQGGYQGEYVDQPKMSEVAQMALEDOATLEYKTIENHSSKSPNGVNLV 871
 758 ABEQPPABEPTPDGLIYARFGRASSHGDFPLTPHASPRRRVV-----SAPFGLDPA 811
 872 ENLDSLP-PKVPQREAST-----GPGASISQCT 898
 812 SAADGLPRPWSPPFTGSLRRPLGPHAPPAATLRR 846

RESULT 7
 US-09-931-836-35
 ; Sequence 35, Application US/09931836
 ; Publication No. US20030027249A1

GENERAL INFORMATION:
 APPLICANT: Desnoyers, Inc
 APPLICANT: Batton, Dan L.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3030R1C1
 CURRENT APPLICATION NUMBER: US/09/931,836
 PRIOR FILING DATE: 2001-08-16
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/112514
 PRIOR FILING DATE: 1998-12-15
 PRIOR APPLICATION NUMBER: 60/113300
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/113430
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/113605
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/113621
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/114140
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/115552
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/116843
 PRIOR FILING DATE: 1999-01-22
 PRIOR APPLICATION NUMBER: 60/125774
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 60/125778
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 60/125826
 PRIOR FILING DATE: 1999-03-24
 PRIOR APPLICATION NUMBER: 60/127035
 PRIOR FILING DATE: 1999-03-31
 PRIOR APPLICATION NUMBER: 60/127706
 PRIOR FILING DATE: 1999-04-05
 PRIOR APPLICATION NUMBER: 60/129122
 PRIOR FILING DATE: 1999-04-13
 PRIOR APPLICATION NUMBER: 60/130359
 PRIOR FILING DATE: 1999-04-21
 PRIOR APPLICATION NUMBER: 60/131270
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131272
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131291
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/132371
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/132379
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/132383
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/135750
 PRIOR FILING DATE: 1999-05-25
 PRIOR APPLICATION NUMBER: 60/138166
 PRIOR FILING DATE: 1999-06-08
 PRIOR APPLICATION NUMBER: 60/144791
 PRIOR FILING DATE: 1999-07-20
 PRIOR APPLICATION NUMBER: 60/146970
 PRIOR FILING DATE: 1999-08-03
 PRIOR APPLICATION NUMBER: 60/162506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 09/311832
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 09/380142

```

PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/516744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908, 827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-23
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 35
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-09-931-836-35

Query Match      36.9%; Score 2009.5; DB 11; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241;

```

```

DB      192  FTAATVDTFLAIDAVIYRSLGDRPTLRVKGHSKMKPEYFVHAIVEMGSHVYFFPREIAE 251
      243  YNTMKVYFPRVAQYCKNDMGSGQVLEKQMTSLFKALNCSVPDSHFEYFVILQAVTV 302
      252  FNYLEKVVSVRAVRCCKNDVGSFVLEKQMTSLFKALNCSVPDSHFEYFVILQAVTV 311
      303  IIRNGRDVLTAFSPYNSIFGSAYCAVDMDLASVFTGRFEKQSPDSTWTPVDERV 362
      312  VSLGRPVYLAVFSPNSIFGSAYCAVDMDLQVAVFBSGRFEKQSPDSTWTPVDERV 371
      363  KRPCCAGSSSLERVATSNFPPDPTLNFIKTHPLMDEAVPSIFNRPFLRMVRYRLTK 422
      372  RRPCCAGAPGV-QYNASSALPDDILNFVKTHPLMDEAVPSIFNRPFLRMVRYRLTK 429
      423  IAVDTAAGRYQNTLVFPGSEKGIILKELAR--IGNSGNLNLSLEMSVYNSEKSYD 480
      430  VAVDVAGAPWGQYLVFPGSENGYLVKFLVRENAVSTGSGISVLEEFETRYRPRDGR 489
      481  GVED--KEIMGQDLDRASSSLVAFSTGVIKYFLRCRCERHGKCKTCIASRDPYGMWE 538
      490  GGGEGQRLSLLEDAAGGLAAPRCVAVRPARCQYSGCMKNCIGSDPYCGMAD 549
      539  GGACSHLSPNSRLTTEODIERGNTDGLDCHNSFVALNGHSSLLPSTTSSTNAEGYE 598
      550  -GSCIFLSPGTRAAFEODVSGASTGLGDC----- 578
      599  SRGMLDWKHLSDSDSTDPPLGAVSHNODKGVIRESYLKGHDQVPTLLAIAVILA 658
      579  -----TGLRLASLSDRGLVSVNLVTSVVA 606
      659  FVMAVFSGITV-YVCVCHRRDVAVQREKE--LTHSRGSMSSVTKLSGLFGDTQSK 715
      607  FVVGAVVSGFSVGMFVGLRERELA--RRKDREALIAGAGVAVLSVSL-----GERRAQ 660
      716  DP-----KPEALITPLMHNGKLAIPGTAK--MLKADQHHLDLTLPTPREST 761
      661  GGGGGGGGGGAGVPPENLAPLMQNG-----WAKTTLQGGPHDSDGLPTPEOT 713
      762  PTLQCKRP-----SRGSEMERQNLIACVDMPMGSPVPTDLP--LRASPEH 811
      714  P-LPQKRLPTPHPHALGPRAMDH-----GHPLLPASASSGILLAPAR 757
      812  IPSVVLPTTQGYGHEVDDQPMSEVAMALEDDAATLEYTITKEHSSKPPNBNVNV 871
      758  AEPQPAPEPTPDGLVYARPARASHGDFPLTPHASPDRRRVV-----SAPTGJLDA 811
      872  ENUDSLP-PKVQREASL-----GPGASISQT 898
      812  SAADGLFRPWSPPPTGSLRRPLGPHAPPAATLRT 846

RESULT 8
US-10-035-977-35
Sequence 35, Application US/10035977
Publication No. US20030134327A1
GENERAL INFORMATION:
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Stewart, Timothy A.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: 23030R1C10
CURRENT APPLICATION NUMBER: US/10/035, 977
CURRENT FILING DATE: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579

```


PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/112514
 PRIOR FILING DATE: 1998-12-15
 PRIOR APPLICATION NUMBER: 60/113300
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/113430
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/113605
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/113621
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/114140
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/115552
 PRIOR FILING DATE: 1998-01-12
 PRIOR APPLICATION NUMBER: 60/116843
 PRIOR FILING DATE: 1999-01-22
 PRIOR APPLICATION NUMBER: 60/125774
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 60/125778
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 60/125826
 PRIOR FILING DATE: 1999-03-24
 PRIOR APPLICATION NUMBER: 60/127035
 PRIOR FILING DATE: 1999-03-31
 PRIOR APPLICATION NUMBER: 60/127706
 PRIOR FILING DATE: 1999-04-05
 PRIOR APPLICATION NUMBER: 60/129122
 PRIOR FILING DATE: 1999-04-13
 PRIOR APPLICATION NUMBER: 60/130359
 PRIOR FILING DATE: 1999-04-21
 PRIOR APPLICATION NUMBER: 60/131270
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131272
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131291
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/132371
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/132379
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/133383
 PRIOR FILING DATE: 1999-05-04
 PRIOR APPLICATION NUMBER: 60/135750
 PRIOR FILING DATE: 1999-05-25
 PRIOR APPLICATION NUMBER: 60/138166
 PRIOR FILING DATE: 1999-06-08
 PRIOR APPLICATION NUMBER: 60/144791
 PRIOR FILING DATE: 1999-07-20
 PRIOR APPLICATION NUMBER: 60/146970
 PRIOR FILING DATE: 1999-08-03
 PRIOR APPLICATION NUMBER: 60/162506
 PRIOR FILING DATE: 1999-10-29
 PRIOR APPLICATION NUMBER: 09/311832
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 09/380142
 PRIOR FILING DATE: 1999-08-25
 PRIOR APPLICATION NUMBER: 09/644848
 PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 09/747259
 PRIOR FILING DATE: 2000-12-20
 PRIOR APPLICATION NUMBER: 09/816744
 PRIOR FILING DATE: 2001-03-22
 PRIOR APPLICATION NUMBER: 09/854208
 PRIOR FILING DATE: 2001-05-10
 PRIOR APPLICATION NUMBER: 09/854280
 PRIOR FILING DATE: 2001-05-10
 PRIOR APPLICATION NUMBER: 09/874503
 PRIOR FILING DATE: 2001-06-05
 PRIOR APPLICATION NUMBER: 09/869599
 PRIOR FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: 09/908,827
 PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: PCT/US99/10733
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: PCT/US99/28551
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/30720
 PRIOR FILING DATE: 1999-12-22
 PRIOR APPLICATION NUMBER: PCT/US00/05601
 PRIOR FILING DATE: 2000-03-01
 PRIOR APPLICATION NUMBER: PCT/US00/05841
 PRIOR FILING DATE: 2000-03-02
 PRIOR APPLICATION NUMBER: PCT/US00/14042
 PRIOR FILING DATE: 2000-05-22
 PRIOR APPLICATION NUMBER: PCT/US00/15264
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: PCT/US00/23522
 PRIOR FILING DATE: 2000-08-23
 PRIOR APPLICATION NUMBER: PCT/US00/23328
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: 2000-12-01
 PRIOR APPLICATION NUMBER: PCT/US00/34956
 PRIOR FILING DATE: 2000-12-20
 PRIOR APPLICATION NUMBER: PCT/US01/06520
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: PCT/US01/17800
 PRIOR FILING DATE: 2001-06-01
 PRIOR APPLICATION NUMBER: PCT/US01/19692
 PRIOR FILING DATE: 2001-06-20
 PRIOR APPLICATION NUMBER: PCT/US01/21066
 PRIOR FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: PCT/US01/21735
 PRIOR FILING DATE: 2001-07-09
 NUMBER OF SEQ ID NOS: 80
 SEQ ID NO 35
 LENGTH: 888
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-035-977-35

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
 Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
 Matches 423; Conservative 130; Mismatches 241;

5 ALLLYFTLLHFAAGFPEDESEPEISISHGNTKYQYVVGHGKFGNTTOR--HRLDIQMIM 62
 12 ALLLLLLLGGAGHLLFFEPPEPEPLSVAPRDYLNHVPVVGSGPGLTBAEGADLNIQGVLL 71
 QY 63 INNGTLYIARDDIYVDTIDTSHTEIYCSKLTWKSROADVDTCRKKGKKEDECHNFIK 122
 Db 72 RVNRTLEIGDRDNIYRVELPEPTSTELRYORKLTWRNSPNDINCRMKKGEGECRNFKV 131
 QY 123 VLLKNDALFVCGTNNFNSCRYYKMDTELEPFGEDESGMARCPYDAKHANVALFADGKL 182
 Db 132 VLLRDESGTLFVCCSNMFPNVCANYSIDTLOPVGDNISGMARCPYDKRANVALFSDGML 191
 QY 183 YSATVTDFLAIDAVIYSLGSEPLRTKVDKSKWLKEFPYQVAVDYGDIYFFFEELAVE 242
 Db 192 FTATVTDFLAIDAVIYSLGSRPFLRTVKHDSKWFKEPYFVHAVERGSHVFFFEELAVE 251
 QY 243 YNTGKTVVPRVAOVCKNDKSGSRVLEKQMTSTLTKARLNSVPGDSHFYNNIIQAVTDV 302
 Db 252 FNYLEKTVVSRVAVCNKNDVGGSPRLVEKQMTSFLKARLNSVPGDSHFYNNIIQAVTGV 311
 QY 303 IIRNGRVVLAIFSTPNYSIPGSAVCAVMDIASVFTGAFKESQSPDSTWTVPYDERVP 362
 Db 312 VSLGRRPVVLAIFSTPNYSIPGSAVCAFDLQVAAVABGRBGRKSPESINTVPYEDQVP 371
 QY 363 KPRPGCCAGSSSLIRYATSNFPPDTNFIKTHPLMDAVSINRPMWFLTMVRYRLTK 422
 Db 372 RPRPGCCAGPM--QYNASSALPDDILNFVYTHPLMDAVSILGHAEMILRTLMRHQLTR 429
 QY 423 IAVTAAQPYQNTHTVVLGSEKGIILKFLAR--IGNSGFLNDSLFLEMSYYNSBKCYD 480

Db 430 VAVDVAGPMGNOTVVFLESEAGTVLKLVRPNASTSGTSLVLEEFETYPDRCGRP 489
 QY 481 GVED--KRMGMQDLBRASSLYVAESTCYIKVPLGCEHNGKCKTKCIASRDYCGMIKE 538
 Db 490 GGGGTGQRLLSTELDLAASGGLLAAPRCVVRVVARCQYSGCMKNCISQDPYCGMAD 549
 QY 539 GGCASHLSPSRLLTFPODIERGNTDGGGCHNSFVALNGHSSSLSTTTSDTAQGYE 598
 Db 550 -GSCIFLSPGTRAFEDVSGASTSLGDC----- 578
 QY 599 SRGMLDMWKHLSDPDSTPLGAVSSHNDKKGVIRESYLKGHDQVPTLLAIVILA 658
 Db 579 -----TGLRASLSEDRAGLVSNILVTSVA 606
 QY 659 FVMGAVFSGITV-YCCDHRKRDVAVVORKEKE--LTHSRGSMSSVTLSGLFGTQSK 715
 Db 607 FVVGAVVSGFSVGMFVGLRERRELA--RRDKKALIAHGADEVLSVSR-----GERRAQ 660
 QY 716 DP-----KREALITPLMHNGKLTATPGNTAK-MLIKADQHILDTALPTPEST 761
 Db 661 GPGRGGGGGGAGVPEALAPLMONG-----WAKATLLGGPHDLSGLLPTPEQT 713
 QY 762 PTLQQRKP-----SRGSEWERNQNLINACTKMPMGSPVITDLP--LRASPSH 811
 Db 714 P-LPQKRLLTPRHHPHALGPRANDH-----GHPDLRASASSSLLLAPAR 757
 QY 812 IBSVVVLPITQOGYQHRYVDPKMSSEVAQMALEDQAITLEYKTIKHLSSKSNHGYNLV 871
 Db 758 APQPAPAPGEPPTDGLLYAARPGRASHGDFPLTPHASPDRRRV-----SAFTGPLDPA 811
 QY 872 ENDSLP-PKVPQREASL-----GPPGASLSQT 898
 Db 812 SAADGLPRPWSPPPTGSLRRPLGPHADPAATLRRT 846

RESULT 9

US-10-137-870-544
 ; Sequence 544, Application US/10137870
 ; Publication No. US2003013883A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroli, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gunney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P33081C15
 ; CURRENT APPLICATION NUMBER: US/10/137,870
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 544
 ; LENGTH: 888
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-137-870-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
 Best Local Similarity 45.2%; Pred. No. 7.4e-160;
 Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLLYFTLLHPAGAGPEDESEPSISHSNNTYKQYEVFGHKRGENTTOR--HRLDIQWIM 62
 Db 12 ALLLLLLLGGHGLPPEPPPLSVAPRDYLNHYEVFGSGPGRLLTPABGADDLNQYV 71
 QY 63 INNGTLYIAARDHITVDIDTSHTEIYCSKLLTKRSQADVDYTCRMKGKHKDECHNPIK 122
 Db 72 RYNRTLLFIQDRNLYKRVLEPPTSTELRYQRLTKRNSPNDINVCRMKGQGECCGNFVX 131
 QY 123 VILKNDALFVCGTNAFNSCRNYKMDTFEPGDFBSGMAFCYDADAKANAVALPADGL 182
 Db 132 VILLDEESTLFVCGSNAPFPCANYSIDTLQVGNISGMARCPYDPKIANALFSDGL 191
 QY 183 YSNATYDPLADAVLYRSLGESPTLRKYHDSKWLKEPVQAVDYGDIYFEFRIAYE 242
 Db 192 FATATYDPLADAVLYRSLGDRPTLRKYHDSKWEPEFVHAVEGSHVYFFREIAYE 251
 QY 243 YNTMGKVPEPPVAVQYCKNDMGSGSVLEKQWTFKARLNCVPGDSHFYFNILQAVTV 302
 Db 252 FMYLEKVVASRYARCKNDVGSFVLEKQWTFKARLNCVPGDSHFYFNILQAVTV 311
 QY 303 IIRNGEDVLLATFSTPPYNSIPGSAYCAVMDLIDASYFTGRFKEQSPDSTWTPVPERP 362
 Db 312 VSLGSPVYLAFFSTPSNSIPGSAYCAFDLTQVAAVFEGRFRDQKSPESIMTPVPEDOVP 371
 QY 363 KRPDGCAGSSSLERATNNEFPDDTINRIKTHPLMDEAVPSIFNRPWFLRTMVRYLRK 422
 Db 372 RRPDGCAGAPGM-QNNASSALPDIDLNVKTHPLMDEAVPSIGHAPWILRTLMRQQLR 429
 QY 423 IAVDTAAGPYQNHVTVFLESEAGTVLKLVRPNASTSGTSLVLEEFETYPDRCGRP 489
 Db 430 VAVDVAGPMGNOTVVFLESEAGTVLKLVRPNASTSGTSLVLEEFETYPDRCGRP 489
 QY 481 GVED--KRMGMQDLBRASSLYVAESTCYIKVPLGCEHNGKCKTKCIASRDYCGMIKE 538
 Db 490 GGGGTGQRLLSTELDLAASGGLLAAPRCVVRVVARCQYSGCMKNCISQDPYCGMAD 549
 QY 539 GGCASHLSPSRLLTFPODIERGNTDGGGCHNSFVALNGHSSSLSTTTSDTAQGYE 598
 Db 550 -GSCIFLSPGTRAFEDVSGASTSLGDC----- 578
 QY 599 SRGMLDMWKHLSDPDSTPLGAVSSHNDKKGVIRESYLKGHDQVPTLLAIVILA 658
 Db 579 -----TGLRASLSEDRAGLVSNILVTSVA 606
 QY 659 FVMGAVFSGITV-YCCDHRKRDVAVVORKEKE--LTHSRGSMSSVTLSGLFGTQSK 715
 Db 607 FVVGAVVSGFSVGMFVGLRERRELA--RRDKKALIAHGADEVLSVSR-----GERRAQ 660
 QY 716 DP-----KREALITPLMHNGKLTATPGNTAK-MLIKADQHILDTALPTPEST 761
 Db 714 P-LPQKRLLTPRHHPHALGPRANDH-----GHPDLRASASSSLLLAPAR 757
 QY 812 IBSVVVLPITQOGYQHRYVDPKMSSEVAQMALEDQAITLEYKTIKHLSSKSNHGYNLV 871
 Db 758 APQPAPAPGEPPTDGLLYAARPGRASHGDFPLTPHASPDRRRV-----SAFTGPLDPA 811
 QY 872 ENDSLP-PKVPQREASL-----GPPGASLSQT 898
 Db 812 SAADGLPRPWSPPPTGSLRRPLGPHADPAATLRRT 846

RESULT 10

US-10-140-018-544
 ; Sequence 544, Application US/10140018
 ; Publication No. US2003013885A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen

```

APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C158
CURRENT APPLICATION NUMBER: US/10/140,018
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-018-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLIYTLHFAAGPEDEPISISGNTKQYPPVGVGKRPNTQR--HRLDIDIM 62
DB 12 ALLIYTLHFAAGPEDEPISISGNTKQYPPVGVGKRPNTQR--HRLDIDIM 71
QY 63 INNGTYIARBDIYVDIDTSHTEIYCSKLTWKSROADVTCRMKGKHCHEHFIK 122
DB 72 RVNRITLFIGRDNLRYELPPTSTELRYQKLTWNSPNDINVCRMKGQEGECRFVK 131
QY 123 VLLKNDALFVCGTNAFNSCRNYKMDLEPFGEDESGMARCEYAKANVALPADGKL 182
DB 132 VLLKNDALFVCGTNAFNSCRNYKMDLEPFGEDESGMARCEYAKANVALPADGKL 191
QY 183 YSAVITFLALDAIYNSLSEPTLRTVKDSKLTKEPPYQAVDYDIIYFFPRELAVE 242
DB 192 FTATVITFLALDAIYNSLSEPTLRTVKDSKLTKEPPYQAVDYDIIYFFPRELAVE 251
QY 243 YNTGKVPFPPVAVGCKNDGSGORVLEKQWTSFLKRLNCSVPGDSHFENILQAVTDV 302
DB 252 FNYLEKVVSRVAVGCKNDGSGORVLEKQWTSFLKRLNCSVPGDSHFENILQAVTDV 311
QY 303 IRNGRDVVLATFSTPYNSIPGSAVCAVMDLDAVFTGRFKECKSPDSTWTPVDERVP 362
DB 312 VSLGGRPVVLAVFSTPNSIPGSAVCAVMDLDAVFTGRFKECKSPDSTWTPVDERVP 371
QY 363 KPRPGCCAGSSSLERYATNSPEPDDTANFKTEPLMDEAVPSIFNRWPLRTWRYRLTK 422
DB 372 RPRPGCCAGSSSLERYATNSPEPDDTANFKTEPLMDEAVPSIFNRWPLRTWRYRLTK 429
QY 423 IAVTAAQPYONHTVVFAGSEKGIILKFLAR--IGNSGFUNDSEFLEEMSVYNEKESYD 480
DB 430 IAVTAAQPYONHTVVFAGSEKGIILKFLAR--IGNSGFUNDSEFLEEMSVYNEKESYD 489
QY 481 GVED--KRMGMQLDRASSSLYYVAFSTCVIKVPLGRCERHCKCKTGIASRDPCYGIKE 538
DB 490 GGETGTGRILSLLEDAASGGLLAFPRCVAVVPVAPRCQYSGCKMCKIGSDPCYGIKE 549
QY 539 GGAASHSPNRSLTFEODIERKNTDGLGDCNNSVALNGHSSLLBTTTSDSTAGSGYE 598
DB 550 -GSCIFLSPGTRAFEDDYSGASTSGLDC-----578
QY 599 SRGMLDMKHLDPDSTDPGLGAVSSHNGDKKGVRESYLGKHDQULVPTLLAIVAVILA 658

```

```

DB 579 -----TGLRLASLSERAGLVSVNLLVTSVAA 606
QY 659 FVMGAVFSGITV-YCYCDHRRKQVAVVQKKE--LTHRRGSMSSVTKLSGLFGDTQSK 715
DB 607 FVMGAVFSGITV-YCYCDHRRKQVAVVQKKE--LTHRRGSMSSVTKLSGLFGDTQSK 715
QY 716 DP-----KREALITPLMNGKLAIPRGNTAK-MLIKADQHHLDTLALPPEST 761
DB 661 GEGRGGGGGGAGVPEPBLALPMONG-----WAKATLLDGGHDDLSGLLPTEQT 713
QY 762 PTLQCRKP-----SRGSRERENQNLINACTKMDPPWGSFVITDUP--LRASPSH 811
DB 714 P-LPQKRLTPPHHPALGPAMDH-----GHPDLASASSSLLLAPAR 757
QY 812 IPSVVLPTTQCYQHEIYDQPKSEVVAQMALEDAQATLEYKITKHLSSKSNHGVNLY 871
DB 758 APEQPPAPGEPPTPDGLYVAPRGASHGDFPLTPHASPERRRVV-----SAPGTLPDA 811
QY 872 ENLDSL-P-KVPGREASL-----GPPGASLSQT 898
DB 812 SAADGLPRPSPPTGSLRRLPGLPHAPPAITLRT 846

RESULT 11
US-10-140-021-544
Sequence 544, Application US/10140021
Publication No. US20030138886A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geriltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C167
CURRENT APPLICATION NUMBER: US/10/140,021
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-021-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLIYTLHFAAGPEDEPISISGNTKQYPPVGVGKRPNTQR--HRLDIDIM 62
DB 12 ALLIYTLHFAAGPEDEPISISGNTKQYPPVGVGKRPNTQR--HRLDIDIM 71
QY 63 INNGTYIARBDIYVDIDTSHTEIYCSKLTWKSROADVTCRMKGKHCHEHFIK 122
DB 72 RVNRITLFIGRDNLRYELPPTSTELRYQKLTWNSPNDINVCRMKGQEGECRFVK 131
QY 123 VLLKNDALFVCGTNAFNSCRNYKMDLEPFGEDESGMARCEYAKANVALPADGKL 182

```

```

Db 132 VLLRDESTLFCGSMANFPVCANYSIDTLQFVGNISGMARCPYDPKHAVALFSDGML 191
QY 183 YSATVTDPLAIDAVIYRSLGESPTRLRTVKHDSKMKKEPYQAVVGYIYFFFEIAYE 242
Db 192 FTAIVTDLAIDAVIYRSLGDRPLRTVKHDSKMKKEPYQAVVGYIYFFFEIAYE 251
QY 243 YNTMKGVVFPVAVQYCKNDMGSGQVLEKQWTSFLKALNCSVPDSSHFFENILQAVTDV 302
Db 252 FVYLEKVVSVARVCKNDVGVSPVLEKQWTSFLKALNCSVPDSSHFFENILQAVTV 311
QY 303 IIRINGDVVLTFTSPYNSIPGSAYCAVMDLDAVFTGFRFEQKSPDSTWTPVDERVP 362
Db 312 VSLGRPVVLAIVFTSPNSIPGSAYCAVMDLQVAVFBERFREQKSPSISWTPVEDQVP 371
QY 363 KRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIENRPFRLRTVRYRLTK 422
Db 372 RRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIENRPFRLRTVRYRLTK 429
QY 423 IAVDTAAGPYQNHVTVFLGSEKGIILKFLAR--IGNSGFLNDSFLFEMSVYNSKCSYD 480
Db 430 VAVDVAGAGWNGQTVFLGSEAGTVLKLVRNASTSGISGLSVLEEFETVRPDRCGRP 489
QY 481 GVED--KRIMGQDLRASSLYVAFSTCYIKVPLRCERHNGCKTKTCLASRDPYCGWIK 538
Db 490 GGGRTGQRLSLDELDAASGGLLAAPRCVAVPVARCQYSGCMKNCIGSDPYCGWAPD 549
QY 539 GGACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSLLPSTTSSTAOEGYE 598
Db 550 -GSCIFLSPGTRAAFEQDVSGASTGLGDC----- 578
QY 599 SRGMDLWKHLDSPESTDPPLGAVSHNQDKKVIRESYLKGHDQVPTVTLAIVILA 658
Db 579 -----TGLRASLSEDRAGLVSNLLVTSVVA 606
QY 659 FVWGAIVSGITV--YCVCDHRKDVAVVQKKE--LTHSRGSMSSVTKLSGLFGDTQSK 715
Db 607 FVWGAIVSGFVGMFGLERRELA--RRKDEKAILAHGAGVAVSVSL--GERRAQ 660
QY 716 DP-----KREAILPLMNGKLAIPGNTAK--MLIKADQHHLDLALPPEST 761
Db 661 GGGGAGGGGGAGVPEALPLMONG-----WAKATLLOGGPHDLSGLLPPEBT 713
QY 762 PTLQQRKP-----SRGSRWERNQNLINACTKMPMGSPVPTDLP--LRASPEH 811
Db 714 P-LPQKRLPLRHPHALGRAWDH-----GHPILPASASSILLALAPAR 757
QY 812 IBSVVVLTPTQCGYQIEYVDQPMSEVAQMALEDQAATLEYKTIKHSKSPNHQVNV 871
Db 758 APQPPAPGEPPTDGLLYAARPGASHODFPPLTPHASPPRRRV-----SAPTPGLDPA 811
QY 872 ENLDSLP--PKVQREASL-----GPPGASLSQT 898
Db 812 SAADGLPRPMSPPPTGSLRRPLGPHAPPAATLART 846

```

RESULT 12

US-10-140-274-544

Sequence 544, Application US/10140274

Publication No. US20030143674A1

GENERAL INFORMATION:

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Denoyere, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.

```

```

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C161
CURRENT APPLICATION NUMBER: US/10/140, 274
PRIORITY FILING DATE: 2002-05-06
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-274-544

```

```

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241;

```

```

QY 5 ALLVFTLLHFAAGFPEDESEPISSHQNTYKQYVFPVGHKRGNTTOR--HRLDIQIM 62
Db 12 ALLVFTLLHFAAGFPEDESEPISSHQNTYKQYVFPVGHKRGNTTOR--HRLDIQIM 71
QY 63 IINGTLYAARDHIYVDITSHTEBETCSKLLTWKSRQADVDTGRMKGKDECHNFIK 122
Db 72 RVNRTLFIGDRDNIYRVLEAPPTSTELRYQKLTWRSNPSIDINVCRMKGQEGECNRFV 131
QY 123 VLLKNDLALFVCGNANFPSCRMVKNMDELPFGDEFGMARCPYDAKHAVALFADGKL 182
Db 132 VLLRDESTLFCGSMANFPVCANYSIDTLQFVGNISGMARCPYDPKHAVALFSDGML 191
QY 183 YSATVTDPLAIDAVIYRSLGESPTRLRTVKHDSKMKKEPYQAVVGYIYFFFEIAYE 242
Db 192 FTAIVTDLAIDAVIYRSLGDRPLRTVKHDSKMKKEPYQAVVGYIYFFFEIAYE 251
QY 243 YNTMKGVVFPVAVQYCKNDMGSGQVLEKQWTSFLKALNCSVPDSSHFFENILQAVTDV 302
Db 252 FVYLEKVVSVARVCKNDVGVSPVLEKQWTSFLKALNCSVPDSSHFFENILQAVTV 311
QY 303 IIRINGDVVLTFTSPYNSIPGSAYCAVMDLDAVFTGFRFEQKSPDSTWTPVDERVP 362
Db 312 VSLGRPVVLAIVFTSPNSIPGSAYCAVMDLQVAVFBERFREQKSPSISWTPVEDQVP 371
QY 363 KRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIENRPFRLRTVRYRLTK 422
Db 372 RRPCCGCAAGM--QYNASSALPDDILNFVKTHPLMDEAVPSIENRPFRLRTVRYRLTK 429
QY 423 IAVDTAAGPYQNHVTVFLGSEKGIILKFLAR--IGNSGFLNDSFLFEMSVYNSKCSYD 480
Db 430 VAVDVAGAGWNGQTVFLGSEAGTVLKLVRNASTSGISGLSVLEEFETVRPDRCGRP 489
QY 481 GVED--KRIMGQDLRASSLYVAFSTCYIKVPLRCERHNGCKTKTCLASRDPYCGWIK 538
Db 490 GGGRTGQRLSLDELDAASGGLLAAPRCVAVPVARCQYSGCMKNCIGSDPYCGWAPD 549
QY 539 GGACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSLLPSTTSSTAOEGYE 598
Db 550 -GSCIFLSPGTRAAFEQDVSGASTGLGDC----- 578
QY 599 SRGMDLWKHLDSPESTDPPLGAVSHNQDKKVIRESYLKGHDQVPTVTLAIVILA 658
Db 579 -----TGLRASLSEDRAGLVSNLLVTSVVA 606
QY 659 FVWGAIVSGITV--YCVCDHRKDVAVVQKKE--LTHSRGSMSSVTKLSGLFGDTQSK 715
Db 607 FVWGAIVSGFVGMFGLERRELA--RRKDEKAILAHGAGVAVSVSL--GERRAQ 660
QY 716 DP-----KREAILPLMNGKLAIPGNTAK--MLIKADQHHLDLALPPEST 761
Db 661 GPGRGGGGGAGVPEALPLMONG-----WAKATLLOGGPHDLSGLLPPEBT 713

```

```

QY 762 PTLQOKRRP-----SRGSEWERNQNLINACTKMPWSPVITPDLR---LRASPSH 811
DB 714 P-LPQKRLPTPHHPHPLALGPRAMDH-----GHPLPLPASASSLLILAPAR 757
QY 812 IPSVYVVLPTIQOQYQHEHYVDQPKKSEVAQVMALEDQAATLEYKITIKELHLSKSPNHGVNLV 871
DB 758 APEQPPAPGEPPTDGRLYAARPGRASHGDFPLTPHASPDRRRVY-----SAPGRPLDPA 811
QY 872 ENDSLPL-PKVPQREASL-----GPPGASLSQT 898
DB 812 SAADGLRPWSPPTGSLRRPLGPHAPPAATLRRT 846

RESULT 13
US-10-140-471-544
; Sequence 544, Application US/10140471
; Publication No. US20030138887A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Defoige, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gunney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C163
; CURRENT APPLICATION NUMBER: US/10/140,471
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 544
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-471-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241;

QY 5 ALLLYFTLLHPRGAGFEPDESEPISSHGNYTKOYFVPGHKGKNTTQR--HRLDIQIM 62
DB 12 ALLLLLLLGGAGHGFPEEPPLSVAPRDYLNHYFVFGSGPRGLTPREGADLINTGRVL 71
QY 63 IANGTLVYAAADHLYTVDIDTSHHEITCSKLTLMKSGQAVDVCQRMKGCKDKCHNPIK 122
DB 72 RVNRTLFIGDRDNDYRVLEPPTSTELRYQRKLTWRSRPSDINVCRMKGCKGEGECRNPFK 131
QY 123 VLLKKNDALFVCGTNAFNPSCRNKMDLLEPFGEFSGMARCEYDAHANVALFADGKL 182
DB 132 VLLRDESTLIVCGSNAPNVCANYISITTLQPVGNISGMARCFDPGHAVVALFSDGML 191
QY 183 YSAIVTDLAIDAVIYRSLGSSPTLRVYKHSKMKIEPYVQAVDGYITFFPREIAVE 242
DB 192 FLATVTDPLAIDAVIYRSLGDRPLTRVYKHSKMKFEPYFVAHVMGSHVYFFFEIAME 251
QY 243 YNTMKVVPFVAVQVCKRDKGSGQVLEKQMTSTLKALINCSVPDGSFYENIIQAVTDV 302
DB 252 FNYLEKVVSVARVAKNDVGGSPVLEKQMTSTLKARLINSVPSDGSFYENVALQAVTG 311
QY 303 IRINGRDVVLATFTSPYNSIPGSAVCAVMDLIASVFTGRFEQKSPDSTWTPVDERVP 362

```

```

DB 312 VSLGPRVVLAVFSTPSNIPGSAVCAFDLTQVAAVEEGFRQKSPESITWTPVEDQVP 371
QY 363 KRPGGCAGSSSLERATNEPDDTLNFKTHPLMDEAVPSIFNRPWFLRTWRYALTK 422
DB 372 RPRPGCCAAFGM--QYNASSALPDDILNFKHPLMDEAVPSIGHAPWILRTLMRHQLTR 429
QY 423 IAVDTAAGPYQHNTVAVFLGSEKGIILKFLAR--IGNSGFNLDSLPLEMSVNSKESYD 480
DB 430 VAVDVAGRWKQNTVAVFLGSEACTVAKFLVRPAASTSGISGLVFLPEFTYPRDRGRP 489
QY 481 GVED--KRIMGCLDRASSGLVYAFSTYKIVPLGRCERHKKCTKTASRDYCGMIKE 538
DB 490 GGGGTGQRLSTLELDASGGLLAAPRCVVRVYVAVRCQYSGCMKNCIGSDPYCGWAPD 549
QY 539 GGCASHLSPNSRLTFPQDIERGNTDGLGCHNSFVALNGHSSSLPSTTSDAQGEYE 598
DB 550 -GSCIFLSPQTRAAFPQDVSGASTSGLDG----- 578
QY 599 SRGGMLDWKHLDSPDSTDLGAVSSHNDKKGVIRESYLKGHDQVPTLLAIVILA 658
DB 579 -----TGLRASLSEDRAGLVSNLVTSSVAA 606
QY 659 FVNGAVFSGITV-YCYCDHRRKDVAVVQKKEK--LTHSRGNSGYTKLSGLFGDTQSK 715
DB 607 FVVGAVVSGFSVGMFVGLRERREILA--RRKDEALLAHGAEVLSYSRL---GERRAQ 660
QY 716 DP-----KREALTLMNGKATGNTAK-MLIRADQHLDTLPTPEST 761
DB 661 GPGRGGGGGGAGVPEALAPLMONG-----WKATLLQGGPHDLSGLPTPEQT 713
QY 762 PTLQOKRRP-----SRGSEWERNQNLINACTKMPWSPVITPDLR---LRASPSH 811
DB 714 P-LPQKRLPTPHHPHPLALGPRAMDH-----GHPLPLPASASSLLILAPAR 757
QY 812 IPSVYVVLPTIQOQYQHEHYVDQPKKSEVAQVMALEDQAATLEYKITIKELHLSKSPNHGVNLV 871
DB 758 APEQPPAPGEPPTDGRLYAARPGRASHGDFPLTPHASPDRRRVY-----SAPGRPLDPA 811
QY 872 ENDSLPL-PKVPQREASL-----GPPGASLSQT 898
DB 812 SAADGLRPWSPPTGSLRRPLGPHAPPAATLRRT 846

RESULT 14
US-10-140-807-544
; Sequence 544, Application US/10140807
; Publication No. US20030134354A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Defoige, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gunney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C174
; CURRENT APPLICATION NUMBER: US/10/140,807
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550

```

SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-922-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

```

5 ALLYFTLLHPAGAPPEDESEPISSHONYTKOYVFGHKRGNTTOR--HRLDIOMIM 62
12 ALLLLLLLGGAGHGFPEEPPLSVAPRDYLNHYVFGSGGRLLTPAEGADDLNIQVLT 71
63 INNGTLYIAARDHIYTVIDTSHTEIYCSKLLTKSKROADVDTORMKXKHECHNF 122
72 RYNRLLFTIGDRDNLRYVELLEPPTSTELRYKRLTWRSNPSDINVRMKKGEGCRNF 131
123 VLLKNDALFVCGTNAFNPSCRNKMDLEPFGEFSGMARCPYDAVGHANVALFADGKL 182
132 VLLRDESTLFVCGSNAPFVPCANYSIDTLQFVGNISGMARCPYDPKHANVALFSDGML 191
183 YSATVTDFLAIDAVIYRSLGSPFLRTYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 242
192 FTAIVTDFLAIDAVIYRSLGDRPTLRITYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 251
243 YNTMGKVYFPRVAQVCKNDMGSSORVLEKOWTSFLKARLNCSPGDSHFYENILQAVTV 302
252 FVYLEKVYVSVARVCKNDVGSSPVLEKOWTSFLKARLNCSPGDSHFYENILQAVTV 311
303 IINGRDVTLATFSTPYNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 362
312 VSLGGRPVVLAIVFSTPSNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 371
363 KRPGCCAGSSSLERATSNPEPDTLNFIKTHPLMDEAVPSIFNRPWLFRTMVRRLTK 422
372 RRPGCCAAPGM--QYNASSALPDDILNVKTHPLMDEAVPSIFNRPWLFRTMVRRLTK 429
423 IAVDTAAGFYONHTVVFLESGEKIILKFLAR--IGNSGLNLSLPLEMSVYNSKCSYD 480
430 VAVDVAGFPGWQTVVFLSGEAGTVLKLVRPNASTSGTSGSVLFEEBETRYRPRDGR 489
481 GVED--KRTMGQDLBRASSLYVASTCYIKVPLRGCEHKGCKTCLASRPYCKMIKE 538
490 GGGETQRLSLLELDASGGLLAAPRCYVAVFVAVACQYSGCMKNCIGSDPYCGWADP 549
539 GGACSHLSPNSRLTFEODIERGTDLGDCNHSFVALNGHSSSLPSTTSDTAQEGYE 598
550 -GSCITLSEGTAAFEQDVSGASTSGLDG----- 578
599 SRGMLDMKHLIDSPTDPLGAVSSHNHODKKGVIRESYLKGHDQLVPTLLAVALIA 658
579 -----TGLRASLSEDRALVAVNLLVTSVAA 606
659 FVWGAFVSGITV-YCYCHRRKQVAVVORKEKE--LTHSRGSMSTVTLGSLGPDTSK 715
607 FVVGAVVSGFVGMFGLDERRELA--RRDKKAILAHAGAEVAVSVSL---GERRKO 660
716 DP-----KPEALITPLMNGKLTATPGNTAK--MLIKADQHHDLTALPTPEST 761
661 GPGRGGGGGGAGVPEALIALPLONG-----WAKATLLGSGHDDSLGLPTPEGT 713
762 PTLQGRKP-----SRGSEVERNQNLINACTKMPNGSVITLDP--LRASPSH 811
714 P-LPQRRLPTPHPHALGPRAMH-----GHPLLPASASSSLLLLAIPAR 757
812 IESVAVLPTQCGYCHYVQPKMSVAVQMALEDQATLEKTIKXHLSSKSPHGVNLY 871
758 APQPAPAGPPTDGGIYAAPRAAGHGFPLTPHASPDRRVV-----SAPLGLDPA 811
872 ENLDSLP-PKVPQREASL-----GPPGASLSQT 898
812 SAADGLPRPWPPTGSLRRRLGPHAPPAATLART 846

```

RESULT 15

US-10-140-922-544
Sequence 544; Application US/10140922
Publication No. US2003013889A1

GENERAL INFORMATION:

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filzaro, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Gettisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACTS ENCODING THE SAME
FILE REFERENCE: P33081C179
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-922-544

```

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7.4e-160; Indels 141; Gaps 20;
Matches 423; Conservative 130; Mismatches 241;

```

5 ALLYFTLLHPAGAPPEDESEPISSHONYTKOYVFGHKRGNTTOR--HRLDIOMIM 62
12 ALLLLLLLGGAGHGFPEEPPLSVAPRDYLNHYVFGSGGRLLTPAEGADDLNIQVLT 71
63 INNGTLYIAARDHIYTVIDTSHTEIYCSKLLTKSKROADVDTORMKXKHECHNF 122
72 RYNRLLFTIGDRDNLRYVELLEPPTSTELRYKRLTWRSNPSDINVRMKKGEGCRNF 131
123 VLLKNDALFVCGTNAFNPSCRNKMDLEPFGEFSGMARCPYDAVGHANVALFADGKL 182
132 VLLRDESTLFVCGSNAPFVPCANYSIDTLQFVGNISGMARCPYDPKHANVALFSDGML 191
183 YSATVTDFLAIDAVIYRSLGSPFLRTYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 242
192 FTAIVTDFLAIDAVIYRSLGDRPTLRITYKHSKWLKEPFYQAVDYGDIYFFFEIAYE 251
243 YNTMGKVYFPRVAQVCKNDMGSSORVLEKOWTSFLKARLNCSPGDSHFYENILQAVTV 302
252 FVYLEKVYVSVARVCKNDVGSSPVLEKOWTSFLKARLNCSPGDSHFYENILQAVTV 311
303 IINGRDVTLATFSTPYNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 362
312 VSLGGRPVVLAIVFSTPSNSIPSSAVCAVMDLDAVFTGRFEQKSPDSTWTPVDEVRP 371
363 KRPGCCAGSSSLERATSNPEPDTLNFIKTHPLMDEAVPSIFNRPWLFRTMVRRLTK 422
372 RRPGCCAAPGM--QYNASSALPDDILNVKTHPLMDEAVPSIFNRPWLFRTMVRRLTK 429
423 IAVDTAAGFYONHTVVFLESGEKIILKFLAR--IGNSGLNLSLPLEMSVYNSKCSYD 480
430 VAVDVAGFPGWQTVVFLSGEAGTVLKLVRPNASTSGTSGSVLFEEBETRYRPRDGR 489

```

QY 481 GVED--KRMGMOLDRASSSLVAFSTCVIKVPLGRCEHGGCKKTCIASRDPYCGWIK 538
 Db 490 GGETGQRLSLLELDNASSGGLAAPPRCVAVPVARCOQYSGCMKNCIGSDPYCGWAPD 549
 QY 539 GGACSHLSPPNSRLTFEODIERGNTDGLDCHNSFVALNGHSSSLLPTTSDSTAQEGYE 598
 Db 550 -GSCIFLSPGTAAFEQDVSGASTGLDC----- 578
 QY 599 SRGGMLDWKHLDSPTDTPLGAVSNHODKGVIRESYLKGHDQVPTLLAIVILA 658
 Db 579 -----TGLRLASLSEDRAGLVSNLVTSSVAA 606
 QY 659 FVMGAVFSGITV-YVCDDRRKDVAVQKKEKE--LTHSRGSMSSVTKLSGLFGDTQSK 715
 Db 607 FVVGAVVSGFVGWTFVGLRERELA--RRKXKALIAHAGAVLSVRL-----GERRAQ 660
 QY 716 DP-----KPEAILTPIMHNGKLATPGNTAK-MLIKADQHHLDLTALPTPEST 761
 Db 661 GPGGRGGGGGAGVPPFALLAPLMQNG-----WAKATLLGGPHDLDSGLLPTPEQT 713
 QY 762 PTLQOKKRP-----SRGSRERENONLINACTDMPPMGSPVPTDLP---LRASPSH 811
 Db 714 P-LPQKRLPTPHPHALGPRAWDH-----GHPILLPASASSSLLLLAPAR 757
 QY 812 IPSVYVLPITQOQYCHEYVDQPKMSEVAAQMALEDQATLEFKTIKEHLSKSPNHGVNLV 871
 Db 758 APPOPPAFGEPTPDGRLYAARPGRASHGDFILTPHASPPRRRV-----SAPTCGLDPA 811
 QY 872 ENLDSLP-PXVPQREASL-----GPPQASLSQT 898
 Db 812 SAADGLPRFMSPPFTGSLRRPLGPHAPPAATLRR 846

Search completed: October 23, 2003, 17:20:47
 Job time : 90 secs


```

QY 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFGEDESGNARCPYDAKHAVALFADG 180
Db 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFGEDESGNARCPYDAKHAVALFADG 180
QY 181 KLYSATVTFPLAIDAVIYRSLGSEPTLRIVKHSKMLKEPFYQAVDYGDYIYFFPREA 240
Db 181 KLYSATVTFPLAIDAVIYRSLGSEPTLRIVKHSKMLKEPFYQAVDYGDYIYFFPREA 240
QY 241 VEYNTMGKVPFPPAVOVCKNDMGSGQRYLEKQWTSFLKRLNCSVGDSHFYFNIIQAYT 300
Db 241 VEYNTMGKVPFPPAVOVCKNDMGSGQRYLEKQWTSFLKRLNCSVGDSHFYFNIIQAYT 300
QY 241 VEYNTMGKVPFPPAVOVCKNDMGSGQRYLEKQWTSFLKRLNCSVGDSHFYFNIIQAYT 300
Db 241 VEYNTMGKVPFPPAVOVCKNDMGSGQRYLEKQWTSFLKRLNCSVGDSHFYFNIIQAYT 300
QY 301 DVIRINGRDVYVLTFTSPYNSIPGSAVCAYMDLJASVFTGRPEKQKSPDSTWTPVDER 360
Db 301 DVIRINGRDVYVLTFTSPYNSIPGSAVCAYMDLJASVFTGRPEKQKSPDSTWTPVDER 360
QY 361 VPKPRPGCCAGSSSLRATSNFPDDTLNFIKTHPLMDEAVPSINRPFILRTWRYRL 420
Db 361 VPKPRPGCCAGSSSLRATSNFPDDTLNFIKTHPLMDEAVPSINRPFILRTWRYRL 420
QY 421 TKIAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGLNDLFLFEMSVYSEKSYD 480
Db 421 TKIAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGLNDLFLFEMSVYSEKSYD 480
QY 481 GVEDKRTMGQDLPRASSLYVAFSTCVIKPLRCRCEHCKCKTCLASRDPIYGMWKEGG 540
Db 481 GVEDKRTMGQDLPRASSLYVAFSTCVIKPLRCRCEHCKCKTCLASRDPIYGMWKEGG 540
QY 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNHSLSLPSITTSSTQOEGYER 600
Db 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNHSLSLPSITTSSTQOEGYER 600
QY 601 GGMLDWGHLDSPDSTPLGAVSHNHQDKGVIRESYLKGHDQVLTLLAVALIAVAF 660
Db 601 GGMLDWGHLDSPDSTPLGAVSHNHQDKGVIRESYLKGHDQVLTLLAVALIAVAF 660
QY 661 MGAVFSGITVYCVCDHRRKQVAVVQKEKELTHSRGSMSSVTKLSGLFGDTSKQKPE 720
Db 661 MGAVFSGITVYCVCDHRRKQVAVVQKEKELTHSRGSMSSVTKLSGLFGDTSKQKPE 720
QY 721 AILTEPLMHNKGLATPGTAKMLIKADQHHLDLTLPTPESTPTLQOKRPSRSGREMER 780
Db 721 AILTEPLMHNKGLATPGTAKMLIKADQHHLDLTLPTPESTPTLQOKRPSRSGREMER 780
QY 781 QNLINACTCKMPGSGSVIPTDPLRASPISHSVVWLPITQOQYQHEVYDQPRMSEVAQ 840
Db 781 QNLINACTCKMPGSGSVIPTDPLRASPISHSVVWLPITQOQYQHEVYDQPRMSEVAQ 840
QY 841 MALEDQAATLEYKTIKESHLSSKSPHNGVNIENIDSLPPRYVPOREASLGPPGASLSQTL 900
Db 841 MALEDQAATLEYKTIKESHLSSKSPHNGVNIENIDSLPPRYVPOREASLGPPGASLSQTL 900
QY 901 SKLEEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNOFGRGNP 960
Db 901 SKLEEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNOFGRGNP 960
QY 961 PAPORVDSICVHSSQPPGQAVTVSRQPSLNAVNSLTRSGLKRTPSLKPDPVPPKSPAPLS 1020
Db 961 PAPORVDSICVHSSQPPGQAVTVSRQPSLNAVNSLTRSGLKRTPSLKPDPVPPKSPAPLS 1020
QY 1021 TSMKPNDACT 1030
Db 1021 TSMKPNDACT 1030

```

```

DE Hypothetical protein KIAA1368 (Fragment).
GN KIAA1368.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL; AB037789; BAA92606.1; -.
DR Genbank; U0010738; SEMA6A.
DR InterPro; IPR003659; Plexin-like.
DR Pfam; PF01403; SEMA; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; SEMA; 1.
KW Hypothetical protein.
FT NON TER.
SQ
SEQUENCE 1049 AA; 11651 MW; 7781D20AC07A8AEA CRC64;
Query Match 99.7%; Score 5431.5; DB 4; Length 1049;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1030; Conservative 0; Mismatches 0; Indels 17; Gaps 1;
QY 1 MRSEALLVFTLLHPAGPPEDESEPTISHSQNYTKQYVGVGKRGKNTQHRIDIQ 60
Db 3 MRSEALLVFTLLHPAGPPEDESEPTISHSQNYTKQYVGVGKRGKNTQHRIDIQ 62
QY 61 IMINGLTYIARDHIYTVDDIDTSHTEEIYCSKLTWKSROADVDTCKMKGKDECHNF 120
Db 63 IMINGLTYIARDHIYTVDDIDTSHTEEIYCSKLTWKSROADVDTCKMKGKDECHNF 122
QY 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFGEDESGNARCPYDAKHAVALFADG 180
Db 121 IKVLLKKNDALFVCGTNAFNPSCRYKMDTLEPFGEDESGNARCPYDAKHAVALFADG 180
QY 181 KLYSATVTFPLAIDAVIYRSLGSEPTLRIVKHSKMLKEPFYQAVDYGDYIYFFPREA 240
Db 181 KLYSATVTFPLAIDAVIYRSLGSEPTLRIVKHSKMLKEPFYQAVDYGDYIYFFPREA 240
QY 241 VEYNTMGKVPFPPAVOVCKNDMGSGQRYLEKQWTSFLKRLNCSVGDSHFYFNIIQAYT 300
Db 241 VEYNTMGKVPFPPAVOVCKNDMGSGQRYLEKQWTSFLKRLNCSVGDSHFYFNIIQAYT 300
QY 301 DVIRINGRDVYVLTFTSPYNSIPGSAVCAYMDLJASVFTGRPEKQKSPDSTWTPVDER 360
Db 301 DVIRINGRDVYVLTFTSPYNSIPGSAVCAYMDLJASVFTGRPEKQKSPDSTWTPVDER 360
QY 361 VPKPRPGCCAGSSSLRATSNFPDDTLNFIKTHPLMDEAVPSINRPFILRTWRYRL 420
Db 361 VPKPRPGCCAGSSSLRATSNFPDDTLNFIKTHPLMDEAVPSINRPFILRTWRYRL 420
QY 421 TKIAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGLNDLFLFEMSVYSEKSYD 480
Db 421 TKIAVDTAAGPYQNHVTVFLGSEKGIILKFLARIGNSGLNDLFLFEMSVYSEKSYD 480
QY 481 GVEDKRTMGQDLPRASSLYVAFSTCVIKPLRCRCEHCKCKTCLASRDPIYGMWKEGG 540
Db 481 GVEDKRTMGQDLPRASSLYVAFSTCVIKPLRCRCEHCKCKTCLASRDPIYGMWKEGG 540
QY 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALN-----GHSSSL 583
Db 541 ACSHLSPNSRLTFEODIERGNTDGLGCHNSFVALNISTPLPDNEMSVYTVYGHSSSL 602
QY 584 PSTTSDSTQOEGYERSGMLDWKHLIDSDSTDPGLGAVSHNHQDKGVIRESYLKGHD 643
Db 603 PSTTSDSTQOEGYERSGMLDWKHLIDSDSTDPGLGAVSHNHQDKGVIRESYLKGHD 662

```

QY 644 QLVPTLLAIAVILAFVWGAVFSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVT 703
 Db 663 QLVPTLLAIAVILAFVWGAVFSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVT 722
 QY 704 KLSGLFGDTQSDKPPREAILPLPMNGKLAIPGNTAKMLIKADQHHLDLTLPTBESTPT 763
 Db 723 KLSGLFGDTQSDKPPREAILPLPMNGKLAIPGNTAKMLIKADQHHLDLTLPTBESTPT 782
 QY 764 LQCKKPSRGSRGSRERNQNLINACTKMDPMPGSPVITPDLPLRASPSHIPSVVVLPIIQQ 823
 Db 783 LQCKKPSRGSRGSRERNQNLINACTKMDPMPGSPVITPDLPLRASPSHIPSVVVLPIIQQ 842
 QY 824 GYQHEVYVQPKKSEVAQWALBDAQATLEKTIKEHLSKSPNHGVLNVENLDSLPKXPQ 883
 Db 843 GYQHEVYVQPKKSEVAQWALBDAQATLEKTIKEHLSKSPNHGVLNVENLDSLPKXPQ 902
 QY 884 REASLGPPGASISQGTLSKRLSEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSN 943
 Db 903 REASLGPPGASISQGTLSKRLSEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSN 962
 QY 944 SSHLSRNSQFGRGNDPPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRT 1003
 Db 963 SSHLSRNSQFGRGNDPPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRT 1022
 QY 1004 PLKPDVPPKPSFAPLSTSMKENDACT 1030
 Db 1023 PLKPDVPPKPSFAPLSTSMKENDACT 1049

RESULT 3

Q9E071 PRELIMINARY; PRT: 1005 AA.

AC Q9E071: 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Axon guidance signal SEMA6A.
 GN SEMA6A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NC NCBITaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=20364339; PubMed=10993894;
 RX Klosternann A., Lutz B., Gertler F., Behl C.;
 RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the disabled/Vasodilator-stimulated Phosphoprotein-1-like protein (EVL) via a novel carboxyl-terminal zyxin-like domain.";
 RL J. Biol. Chem. 275:39647-39653(2000).
 DR EMBL/AF288666; AGS29494.1; -;
 DR MGD; MGI:1203727; Sema6a.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PST; 1.
 DR SMART; SM00630; Sema; 1.
 SQ SEQUENCE 1005 AA; 111758 MW; 57B6927F45B079D CRC64;

Query Match 92.2%; Score 5024.5; DB 11; Length 1005;

Best Local Similarity 91.9%; Pred. No. 0;
 Matches 948; Conservative 26; Mismatches 30; Indels 27; Gaps 2;

QY 1 MRSEALLVYTLIFPACAGFPEDESEPTISGHANYTYQYPFVGHKGRNTTQHRRLDIQ 60
 Db 1 MRDAALLCTLLHAGAGFPEDESEPTISGHANYTYQYPFVGHKGRNTTQHRRLDIQ 60
 QY 61 IMTNGTLYAARDHITYVDITDSSHTEIYCSKKITMKSQAQVDCRMKGKHKDCGNF 120
 Db 61 IMTNRLLVVAARDHITYVDITDSSHTEIYCSKKITMKSQAQVDCRMKGKHKDCGNF 120

QY 121 IKVLLKNDALFVCGTNAFNPSCRNRYKMDTLPEFGDEFSGMARCPYDAKHANVALFPADG 180
 Db 121 IKVLLKNDALFVCGTNAFNPSCRNRYKMDTLPEFGDEFSGMARCPYDAKHANVALFPADG 180
 QY 181 KLYGATYDPLAIDAVYRSLSGESPTLRVYHDSKWLKEPFYQAVDYGDYIYFFFRISA 240
 Db 181 KLYGATYDPLAIDAVYRSLSGESPTLRVYHDSKWLKEPFYQAVDYGDYIYFFFRISA 240
 QY 241 VEYNTMGKVPFPAVAYQCKDMGSGQRYLEKQWTSFLKARLNCVPGDSEHYFNILQAVT 300
 Db 241 VEYNTMGKVPFPAVAYQCKDMGSGQRYLEKQWTSFLKARLNCVPGDSEHYFNILQAVT 300
 QY 301 DVIRINGRDVTLAFSTPPYNSIPGSAVAYMDLASFTRGRFQKSPDSTWTVPDER 360
 Db 301 DVIRINGRDVTLAFSTPPYNSIPGSAVAYMDLASFTRGRFQKSPDSTWTVPDER 360
 QY 361 VPKRPGCCAGSSSLERATSNPEPDDPLNFKTHPLMDEAVPSIFNSPMFLRTWRYRL 420
 Db 361 VPKRPGCCAGSSSLERATSNPEPDDPLNFKTHPLMDEAVPSIFNSPMFLRTWRYRL 420
 QY 421 TKIADTAAGFYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFLEEMSVYNSKCYD 480
 Db 421 TKIADTAAGFYQNHVTVFLGSEKGIILKFLARIGNSGFLNDSLFLEEMSVYNSKCYD 480
 QY 481 GVEDKRIKGMQLDRASSSLYYAFSTCVIKVPLGRCHRGKCKTCIASRDPYCGWIKXG 540
 Db 481 GVEDKRIKGMQLDRASSSLYYAFSTCVIKVPLGRCHRGKCKTCIASRDPYCGWIKXG 540
 QY 541 ACSHLSFNSRLTFBODIERGNTDGLGDCNHSFVALNGHSSILPSTTSDS7AOCGYSR 600
 Db 541 ACSHLSFNSRLTFBODIERGNTDGLGDCNHSFVALNGHSSILPSTTSDS7AOCGYSR 600
 QY 575 SCALHSPSRDLTFBODIERGNTDGLGDCNHSFVALNGHSSILPSTTSDS7AOCGYSR 634
 Db 575 SCALHSPSRDLTFBODIERGNTDGLGDCNHSFVALNGHSSILPSTTSDS7AOCGYSR 634
 QY 601 GMDLMDKHLDSPOSTDPLGAVSSHNHODKKVIRESYLKAGHDQVPTLLAIAVILAFV 660
 Db 601 GMDLMDKHLDSPOSTDPLGAVSSHNHODKKVIRESYLKAGHDQVPTLLAIAVILAFV 660
 QY 661 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDKPE 720
 Db 661 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDKPE 720
 QY 694 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDKPE 764
 Db 694 MGAVVSGITVYCVCDHRKRDVAVYQKKEKELTHSRGSMSSVTXLSGFGDTQSDKPE 764
 QY 721 AILPLMNGKLAIPGNTAKMLIKADQHHLDLTLPTBESTPTLQCKKPSRGSRERN 780
 Db 721 AILPLMNGKLAIPGNTAKMLIKADQHHLDLTLPTBESTPTLQCKKPSRGSRERN 780
 QY 754 AILPLMNGKLAIPGNTAKMLIKADQHHLDLTLPTBESTPTLQCKKPSRGSRERN 814
 Db 754 AILPLMNGKLAIPGNTAKMLIKADQHHLDLTLPTBESTPTLQCKKPSRGSRERN 814
 QY 781 QNLINACTKMDPMPGSPVITPDLPLRASPSHIPSVVVLPIIQQGYQHEVYVDPKXSE-VA 839
 Db 781 QNLINACTKMDPMPGSPVITPDLPLRASPSHIPSVVVLPIIQQGYQHEVYVDPKXSE-VA 839
 QY 840 QNALBDAQATLEKTIKEHLSKSPNHGVLNVENLDSLPKYPQREASLGPPGASISQGT 899
 Db 840 QNALBDAQATLEKTIKEHLSKSPNHGVLNVENLDSLPKYPQREASLGPPGASISQGT 899
 QY 874 QNALBDAQATLEKTIKEHLSKSPNHGVLNVENLDSLPKYPQREASLGPPGASISQGT 934
 Db 874 QNALBDAQATLEKTIKEHLSKSPNHGVLNVENLDSLPKYPQREASLGPPGASISQGT 934
 QY 900 LSKRLSEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 959
 Db 900 LSKRLSEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 959
 QY 934 LSKRLSEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 994
 Db 934 LSKRLSEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQFGRGNDP 994
 QY 960 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKDVPPKPSFAPL 1019
 Db 960 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKDVPPKPSFAPL 1019
 QY 994 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKDVPPKPSFAPL 1059
 Db 994 PPAPQVDSIQVHSSQPSGOAVTVSRQPSLAINVNSLTRSGLKRTSLKDVPPKPSFAPL 1059

RESULT 4

Q96SW4 PRELIMINARY; PRT: 699 AA.

ID Q96SW4: 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Hypothetical protein Flj14595.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Niimiya K., Iwayanagi T.,
 RT "NEO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027501; BAB55158.1;
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 699 AA; 76723 MW; 2E5F11D59741394 CRC64;
 Query Match 67.5%; Score 3677; DB 4; Length 699;
 Best Local Similarity 100.0%; Pred. No. 1.5e-290; Indels 0; Gaps 0;
 Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 332 MLDIASVFTGRFKEOKSPDSTWTPVPPDERVPPKPPGCGAGSSSILERYATSNEFPDPTLNF 391
 DB 1 MLDIASVFTGRFKEOKSPDSTWTPVPPDERVPPKPPGCGAGSSSILERYATSNEFPDPTLNF 60
 QY 392 IKTPLMDEAVPSIFNRPWFLRTWRYRLTKIADVTAAQYQNTVPLGSEKGIILKFL 451
 DB 61 IKTPLMDEAVPSIFNRPWFLRTWRYRLTKIADVTAAQYQNTVPLGSEKGIILKFL 120
 QY 452 ARGNSEFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQLDRASSSLVYAFSTCYIKVP 511
 DB 121 ARGNSEFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQLDRASSSLVYAFSTCYIKVP 180
 QY 512 LGRCEHKGCKKCTCIASRDPCYCGWIKEGACSHLSPNSRLTFEODIERGNTDGLGDCNS 571
 DB 181 LGRCEHKGCKKCTCIASRDPCYCGWIKEGACSHLSPNSRLTFEODIERGNTDGLGDCNS 240
 QY 572 FVALNGSSSLPSTTSDTAQEGYSGRGMLDWKHLIDSPDSTPLGAVSSHNDKK 631
 DB 241 FVALNGSSSLPSTTSDTAQEGYSGRGMLDWKHLIDSPDSTPLGAVSSHNDKK 300
 QY 632 GVIESYLKGHQCVPTLLAIVLAFVWGA VFSGITVYCVCDHRKRDVAVYQREKEL 691
 DB 301 GVIESYLKGHQCVPTLLAIVLAFVWGA VFSGITVYCVCDHRKRDVAVYQREKEL 360
 QY 692 THSRGSMSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLLD 751
 DB 361 THSRGSMSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLLD 420
 QY 752 LTAIPTEBSPTLLOQRKPSRGSEWERNQNLINACTKMPMGSPVITPDLPLRASP 811
 DB 421 LTAIPTEBSPTLLOQRKPSRGSEWERNQNLINACTKMPMGSPVITPDLPLRASP 480
 QY 812 IPSVYVLPITQOGYQHEVYDQPKMSVYQALDQAATLEKTIKEHLSSKSPNHGVLV 871
 DB 481 IPSVYVLPITQOGYQHEVYDQPKMSVYQALDQAATLEKTIKEHLSSKSPNHGVLV 540
 QY 872 ENLDSLPRKVPQREASIGPAGASISQTLGSKRLEMHSSSYGVYKRSYPTNSLTRSHQA 931
 DB 541 ENLDSLPRKVPQREASIGPAGASISQTLGSKRLEMHSSSYGVYKRSYPTNSLTRSHQA 600
 QY 932 TTLKRNNTSSNSHLSRNSQFGRGNDPPAPORVDSIQVHSSQPSGQAVTVSRQPSLNA 991
 DB 601 TTLKRNNTSSNSHLSRNSQFGRGNDPPAPORVDSIQVHSSQPSGQAVTVSRQPSLNA 660
 QY 992 YNSLTRSGIKRTPSLKPDVPPKPSFAPLSTSMKPNDACT 1030

DB 661 YNSLTRSGIKRTPSLKPDVPPKPSFAPLSTSMKPNDACT 699
 RESULT 5
 ID Q96SM8 PRELIMINARY; PRT; 574 AA.
 AC Q96SM8;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Hypothetical protein F141478.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Niimiya K., Iwayanagi T.,
 RT "NEO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027501; BAB55158.1;
 DR InterPro: IPR003659; Plexin-like.
 DR SMART; SM00423; PSI; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 574 AA; 62822 MW; 0C79E01A4117A495 CRC64;
 Query Match 54.7%; Score 2980.5; DB 4; Length 574;
 Best Local Similarity 90.7%; Pred. No. 6.6e-234; Indels 59; Gaps 1;
 Matches 574; Conservative 0; Mismatches 0; Indels 59; Gaps 1;
 QY 398 MDEAVPSIFNRPWFLRTWRYRLTKIADVTAAQYQNTVPLGSEKGIILKFLARIGNS 457
 DB 1 MDEAVPSIFNRPWFLRTWRYRLTKIADVTAAQYQNTVPLGSEKGIILKFLARIGNS 20
 QY 458 GFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQLDRASSSLVYAFSTCYIKVP 517
 DB 21 GFLNDSLFLEMSVYNSEKSYDGEDKRIKMGQLDRASSSLVYAFSTCYIKVP 61
 QY 518 HGCKKCTCIASRDPCYCGWIKEGACSHLSPNSRLTFEODIERGNTDGLGDCNSFYALNG 577
 DB 62 HGCKKCTCIASRDPCYCGWIKEGACSHLSPNSRLTFEODIERGNTDGLGDCNSFYALNG 121
 QY 578 HSSSLPSTTSDTAQEGYSGRGMLDWKHLIDSPDSTPLGAVSSHNDKKGIRES 637
 DB 122 HSSSLPSTTSDTAQEGYSGRGMLDWKHLIDSPDSTPLGAVSSHNDKKGIRES 181
 QY 638 YLKGHQLVPTLLAIVLAFVWGA VFSGITVYCVCDHRKRDVAVYQREKELTHSRG 697
 DB 182 YLKGHQLVPTLLAIVLAFVWGA VFSGITVYCVCDHRKRDVAVYQREKELTHSRG 241
 QY 698 SMSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLLD 757
 DB 242 SMSSVTKLSLFGDTOSKDPKPEAILPLMHNGKLTATGNTAKMLIKADQHLLD 301
 QY 758 PESTPTLLOQRKPSRGSEWERNQNLINACTKMPMGSPVITPDLPLRASP 817
 DB 302 PESTPTLLOQRKPSRGSEWERNQNLINACTKMPMGSPVITPDLPLRASP 361
 QY 818 LPTQOGYQHEVYDQPKMSVYQALDQAATLEKTIKEHLSSKSPNHGVLV 877
 DB 362 LPTQOGYQHEVYDQPKMSVYQALDQAATLEKTIKEHLSSKSPNHGVLV 421
 QY 878 PPKVPQREASIGPAGASISQTLGSKRLEMHSSSYGVYKRSYPTNSLTRSHQA 937
 DB 422 PPKVPQREASIGPAGASISQTLGSKRLEMHSSSYGVYKRSYPTNSLTRSHQA 481
 QY 938 NTSSNSHLSRNSQFGRGNDPPAPORVDSIQVHSSQPSGQAVTVSRQPSLNA 997

Db 482 NTSSNSHLSRNSPFGNDPPAPFQVDSIGVHSSQPGQVTVSRQSLNAYNSLTR 541
QY 998 SGLKRTPLSKDPVPPKPSFAPLSTMKPNDACT 1030
Db 542 SGLKRTPLSKDPVPPKPSFAPLSTMKPNDACT 574
RESULT 6
Q8BUT0 PRELIMINARY; PRT; 587 AA.
ID Q8BUT0
AC Q8BUT0
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Sema domain.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK082711; BAC38582.1;
SO SEQUENCE 587 AA; 66044 MW; EB9C7B102C4DB97A CRC64;
Query Match 54.4%; Score 2963; DB 11; Length 587;
Best local Similarity 95.1%; Pred. No. 1.8e-232;
Matches 549; Conservative 13; Mismatches 15; Indels 0; Gaps 0;
QY 1 MRSEALLVFTLLHFAAGFPEDESEPISSHGNTYKQPVFVGHKPGKNTTQRRLDIOM 60
Db 1 MRPALLLCTLLHCAAGFPEDESEPISSHGNTYKQPVFVGHKPGKNTTQRRLDIOM 60
QY 61 IMINAGTLYIARADHITVDIDTSHTEBIYCSKLTWKSROADVTCRMKGKDECHNF 120
Db 61 IMINRRTLYVARADHITVDIDTSHTEBIYCSKLTWKSROADVTCRMKGKDECHNF 120
QY 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
Db 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
QY 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
Db 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
QY 181 KLYSATVDFLAIDAVIYRSJGDSPTLRIVKXDSKMLKEPYFVQAVDYGDIYFFFRRIA 240
Db 181 KLYSATVDFLAIDAVIYRSJGDSPTLRIVKXDSKMLKEPYFVQAVDYGDIYFFFRRIA 240
QY 241 VEYNTMGKVPFPRVAQVCKNDMGSSQRYLEKQWTSFLKARLNCSPGDSHFYFNILQAVT 300
Db 241 VEYNTMGKVPFPRVAQVCKNDMGSSQRYLEKQWTSFLKARLNCSPGDSHFYFNILQAVT 300
QY 301 DVIRINGRDVVLATFSTPYNSIPGSAYCAVMDLIASVFTGRFKEQKSPDSTWTFVPPER 360
Db 301 DVIRINGRDVVLATFSTPYNSIPGSAYCAVMDLIASVFTGRFKEQKSPDSTWTFVPPER 360
QY 361 VPKPRPGCCAGSSSLEKATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWFLRTWVRYL 420
Db 361 VPKPRPGCCAGSSSLEKATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWFLRTWVRYL 420
QY 421 TKIADVTAAGPYQNTTVVFLSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSKCSYD 480
Db 421 TKIADVTAAGPYQNTTVVFLSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSKCSYD 480
QY 481 GVEDKRIMGQULDRASSSLIYAFSTCVIKVPLGRCEBHGCKKCTIASRDPYCGWVRESG 540
Db 481 GVEDKRIMGQULDRASSSLIYAFSTCVIKVPLGRCEBHGCKKCTIASRDPYCGWVRESG 540
QY 541 ACSHSPNSRLTFEODIERGNTDGLGDCNSFVALNG 577

Db 541 SCAHLSPLSLTFEODIERGNTDGLGDCNSFVALNG 577
RESULT 7
Q8BXZ7 PRELIMINARY; PRT; 605 AA.
ID Q8BXZ7
AC Q8BXZ7
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Sema domain.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK042751; BAC31351.1;
SO SEQUENCE 605 AA; 68299 MW; B4FE713BE8AF6E90 CRC64;
Query Match 54.3%; Score 2961; DB 11; Length 605;
Best local Similarity 94.5%; Pred. No. 2.8e-212;
Matches 550; Conservative 14; Mismatches 18; Indels 0; Gaps 0;
QY 1 MRSEALLVFTLLHFAAGFPEDESEPISSHGNTYKQPVFVGHKPGKNTTQRRLDIOM 60
Db 1 MRPALLLCTLLHCAAGFPEDESEPISSHGNTYKQPVFVGHKPGKNTTQRRLDIOM 60
QY 61 IMINAGTLYIARADHITVDIDTSHTEBIYCSKLTWKSROADVTCRMKGKDECHNF 120
Db 61 IMINRRTLYVARADHITVDIDTSHTEBIYCSKLTWKSROADVTCRMKGKDECHNF 120
QY 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
Db 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
QY 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
Db 121 IKVLLKNDALFVCGTNAFNPSCNRYKMDLTFEPGDFSGMARCPYDAKANAALFADG 180
QY 181 KLYSATVDFLAIDAVIYRSJGDSPTLRIVKXDSKMLKEPYFVQAVDYGDIYFFFRRIA 240
Db 181 KLYSATVDFLAIDAVIYRSJGDSPTLRIVKXDSKMLKEPYFVQAVDYGDIYFFFRRIA 240
QY 241 VEYNTMGKVPFPRVAQVCKNDMGSSQRYLEKQWTSFLKARLNCSPGDSHFYFNILQAVT 300
Db 241 VEYNTMGKVPFPRVAQVCKNDMGSSQRYLEKQWTSFLKARLNCSPGDSHFYFNILQAVT 300
QY 301 DVIRINGRDVVLATFSTPYNSIPGSAYCAVMDLIASVFTGRFKEQKSPDSTWTFVPPER 360
Db 301 DVIRINGRDVVLATFSTPYNSIPGSAYCAVMDLIASVFTGRFKEQKSPDSTWTFVPPER 360
QY 361 VPKPRPGCCAGSSSLEKATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWFLRTWVRYL 420
Db 361 VPKPRPGCCAGSSSLEKATSNPEPDDTLNFKTHPLMDEAVPSIFNRPWFLRTWVRYL 420
QY 421 TKIADVTAAGPYQNTTVVFLSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSKCSYD 480
Db 421 TKIADVTAAGPYQNTTVVFLSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSKCSYD 480
QY 481 GVEDKRIMGQULDRASSSLIYAFSTCVIKVPLGRCEBHGCKKCTIASRDPYCGWVRESG 540
Db 481 GVEDKRIMGQULDRASSSLIYAFSTCVIKVPLGRCEBHGCKKCTIASRDPYCGWVRESG 540
QY 541 ACSHSPNSRLTFEODIERGNTDGLGDCNSFVALNGSSSL 582
Db 541 SCAHLSPLSLTFEODIERGNTDGLGDCNSFVALNDISTPL 582

RESULT 8

096SY4 PRELIMINARY; PRT: 562 AA.
 ID 096SY4
 AC 096SY4
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Hypothetical protein FLN14565.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 Isogai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Matanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Makematsu A., Nakamura Y., Nagahara K., Masuno Y.,
 RA Nimomiya K., Iwayanagi T.,
 RT "NEBO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK027471; BAB55136.1;
 DR InterPro: IPR003659; Plexin-like.
 DR SMART; SM00423; PSI; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 562 AA; 61313 MW; 6A83685FAD1DD78A CRC64;

Query Match 54.2%; Score 2952; DB 4; Length 562;
 Best Local Similarity 100.0%; Pred. No. 1.3e-231;
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 MSVYNEKCSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGGCKKTCIAS 528
 DB 1 MSVYNEKCSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGGCKKTCIAS 60
 QY 529 RDPYCGWIKREGAGCSHLSFNSRLTFEODIERGNTDGLGCHNSFYALNHSSSLPSTT 588
 DB 61 RDPYCGWIKREGAGCSHLSFNSRLTFEODIERGNTDGLGCHNSFYALNHSSSLPSTT 120
 QY 589 SDSTAQEGYESRGMLDWKHLSDPSDTPPLGAVSSHNDKKGVIRESYLKGHDQLVYV 648
 DB 121 SDSTAQEGYESRGMLDWKHLSDPSDTPPLGAVSSHNDKKGVIRESYLKGHDQLVYV 180
 QY 649 TLLAIAVILAFVWGAVFSGITVYCVCDHRRKDVAVVQRKEKELTSSRSGMSVTKLSGL 708
 DB 181 TLLAIAVILAFVWGAVFSGITVYCVCDHRRKDVAVVQRKEKELTSSRSGMSVTKLSGL 240
 QY 709 FGDTSKDPKPEALITPLMHNGKLATPGNTAKMLIKADQHLDLTLPTRESTPTLQOKR 768
 DB 241 FGDTSKDPKPEALITPLMHNGKLATPGNTAKMLIKADQHLDLTLPTRESTPTLQOKR 300
 QY 769 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPISHVYVLPITQOQYQHE 828
 DB 301 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPISHVYVLPITQOQYQHE 360
 QY 829 YVDQPKMSVQAQALDQAALEKYTKIKELSSKSPHNGVNVLENLDSLPKVPQREASL 888
 DB 361 YVDQPKMSVQAQALDQAALEKYTKIKELSSKSPHNGVNVLENLDSLPKVPQREASL 420
 QY 889 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 948
 DB 421 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 480
 QY 949 RNQSPGRGNDPPAPQVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPLK 1008
 DB 481 RNQSPGRGNDPPAPQVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPLK 540
 QY 1009 DVPPKPSFAPLSTSMKENDACT 1030
 DB 541 DVPPKPSFAPLSTSMKENDACT 562

RESULT 9

08NC49 PRELIMINARY; PRT: 562 AA.
 ID 08NC49
 AC 08NC49
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Hypothetical protein FLJ90494.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 Isogai T., Ota T., Nishikawa T., Hayaashi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai H., Saito K.,
 RA Yamamoto J., Makematsu A., Nakamura Y., Kojima S., Nagahara K.,
 RA Masuno Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Nimomiya K.,
 RT "NEBO human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK074975; BAC11326.1;
 DR InterPro: IPR003659; Plexin-like.
 DR SMART; SM00423; PSI; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 562 AA; 61286 MW; 708041459B34D78A CRC64;

Query Match 54.1%; Score 2947; DB 4; Length 562;
 Best Local Similarity 99.8%; Pred. No. 3.4e-231;
 Matches 561; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 469 MSVYNEKCSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGGCKKTCIAS 528
 DB 1 MSVYNEKCSYDGVEDKRMGMQLDRASSSLVVAESTCYIKYPLGRCERHGGCKKTCIAS 60
 QY 529 RDPYCGWIKREGAGCSHLSFNSRLTFEODIERGNTDGLGCHNSFYALNHSSSLPSTT 588
 DB 61 RDPYCGWIKREGAGCSHLSFNSRLTFEODIERGNTDGLGCHNSFYALNHSSSLPSTT 120
 QY 589 SDSTAQEGYESRGMLDWKHLSDPSDTPPLGAVSSHNDKKGVIRESYLKGHDQLVYV 648
 DB 121 SDSTAQEGYESRGMLDWKHLSDPSDTPPLGAVSSHNDKKGVIRESYLKGHDQLVYV 180
 QY 649 TLLAIAVILAFVWGAVFSGITVYCVCDHRRKDVAVVQRKEKELTSSRSGMSVTKLSGL 708
 DB 181 TLLAIAVILAFVWGAVFSGITVYCVCDHRRKDVAVVQRKEKELTSSRSGMSVTKLSGL 240
 QY 709 FGDTSKDPKPEALITPLMHNGKLATPGNTAKMLIKADQHLDLTLPTRESTPTLQOKR 768
 DB 241 FGDTSKDPKPEALITPLMHNGKLATPGNTAKMLIKADQHLDLTLPTRESTPTLQOKR 300
 QY 769 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPISHVYVLPITQOQYQHE 828
 DB 301 KPSRSGREWERNQNLINACTKMPMGSPVITPDLPLRASPISHVYVLPITQOQYQHE 360
 QY 829 YVDQPKMSVQAQALDQAALEKYTKIKELSSKSPHNGVNVLENLDSLPKVPQREASL 888
 DB 361 YVDQPKMSVQAQALDQAALEKYTKIKELSSKSPHNGVNVLENLDSLPKVPQREASL 420
 QY 889 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 948
 DB 421 GPPGASLSQGLSKRLEMHSSSYGVYKRSYPTNSLTRSHQATTLKRNNTNSNSHLS 480
 QY 949 RNQSPGRGNDPPAPQVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPLK 1008
 DB 481 RNQSPGRGNDPPAPQVDSIQVHSSQPSGQAVTVSRQPSLAINVSLTRSGLKRTPLK 540
 QY 1009 DVPPKPSFAPLSTSMKENDACT 1030
 DB 541 DVPPKPSFAPLSTSMKENDACT 562

RESULT 10

096T04

ID Q96T04 PRELIMINARY; PRT; 507 AA.
 AC Q96T04:
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ14533.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Iisogi T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Nakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
 RA Niimura K., Iwayanagi T.;
 RA "NEO human cDNA sequencing project";
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AK027439; BAB511.1;
 DR InterPro: IPR003659; Pfam: PF01403; SMART: SM00423; PSI; 1.
 DR SMART: SM00423; PSI; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 507 AA; 55464 MW; 8CC567B43BC51B3 CRC64;

Query Match 48.1%; Score 2621.5; DB 4; Length 507;
 Best Local Similarity 90.2%; Pred. No. 9.8e-205;
 Matches 507; Conservative 0; Mismatches 0; Indels 55; Gaps 1;
 DB 469 MSVNSEKSYGVVDKRMQGLDRASSSLVAFSTCVIKYPLRCERHGGCKTCTAS 528
 1 MSVNSEKSYGVVDKRMQGLDRASSSLVAFSTCVIKYPLRCERHGGCKTCTAS 60
 QY 529 RDPYGMKIEGACGCHLSPNSRLTFEODIERGNTGLGDCNHSFYALNGHSSLLPSTT 588
 61 RDPYGMKIEGACGCHLSPNSRLTFEODIERGNTGLGDCNHSFYALNGHSSLLPSTT 108
 DB 589 SDSTAQEGYESRGMLDWKHLDPDSTDPGLGAVSSHNDKKGYIRRESYLKGHDQYLV 648
 109 -----GVIRRESYLKGHDQYLV 125
 DB 649 TLTAIAVILAFYMGAVFSGITVYCVCDHRRKQAVVQREKELTSSRRSGMSVTKLSGL 708
 126 TLTAIAVILAFYMGAVFSGITVYCVCDHRRKQAVVQREKELTSSRRSGMSVTKLSGL 185
 QY 709 FGDTSQDKPKPAIILPLMHNGKLTAPGNTAKMLIKADQHLDLTLPTPESTPTLOQR 768
 186 FGDTSQDKPKPAIILPLMHNGKLTAPGNTAKMLIKADQHLDLTLPTPESTPTLOQR 245
 DB 769 KPSRSGREWERQNINACTKMPMGSVIPTDLPLRASPSPHIBPVVTLPTTQOQYHE 828
 246 KPSRSGREWERQNINACTKMPMGSVIPTDLPLRASPSPHIBPVVTLPTTQOQYHE 305
 QY 829 YNDQPKSVYAQMALDAQATLEKTIKHEHLSKSNHGVNLYENLDSLPKYPQREAST 888
 306 YNDQPKSVYAQMALDAQATLEKTIKHEHLSKSNHGVNLYENLDSLPKYPQREAST 365
 DB 889 GPPGASLQSGLSKRLMEHSSSYGVYKRSYPTNSLTRSHQATLLKRNNTSSNSHLS 948
 366 GPPGASLQSGLSKRLMEHSSSYGVYKRSYPTNSLTRSHQATLLKRNNTSSNSHLS 425
 QY 949 RNQSPFRGNDPPAPFQVDSIQVHSSQPSGQAVTVRQPSLNAYNLSITSGKRTSLKP 1008
 426 RNQSPFRGNDPPAPFQVDSIQVHSSQPSGQAVTVRQPSLNAYNLSITSGKRTSLKP 485
 DB 1009 DVPKPSFAPLSTSMKNDACT 1030
 486 DVPKPSFAPLSTSMKNDACT 507
 RESULT 11

Q98FY4 PRELIMINARY; PRT; 1073 AA.
 AC Q98FY4:
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Semaphorin 6D isoform 4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Qu X., Zhai Y., Wei H., Yu Y., Tang F., He F.;
 RA "Homo sapiens semaphorin 6D isoform 4 (SEMA6D.4) mRNA, complete cds";
 RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF389429; AAM69452.1;
 DR InterPro: IPR001627; Pfam: PF01403; SMART: SM00630; SEMA; 1.
 DR SMART: SM00630; SEMA; 1.
 SQ SEQUENCE 1073 AA; 119872 MW; 7DCB4DFC5BF70F9E CRC64;

Query Match 43.4%; Score 2368; DB 4; Length 1073;
 Best Local Similarity 46.0%; Pred. No. 1.5e-183;
 Matches 512; Conservative 159; Mismatches 308; Indels 134; Gaps 30;
 DB 1 MRSEALLIYFTLL--HFAGAPPEDEPPTISHGNTKQYVFGHPRGRTTQRHLD 57
 1 MRSEALLIYFTLL--HFAGAPPEDEPPTISHGNTKQYVFGHPRGRTTQRHLD 58
 QY 58 IOMINMGTYIARHDIYVDIDTSHTEIEYCSKLTWRSRQADVDTGRMKGHKDEC 117
 59 FOLMKIKDTIYIARHDIYVDIDTSHTEIEYCSKLTWRSRQADVDTGRMKGHKDEC 118
 DB 118 HNFIVKLLKNDDALFVCGTNAFNSCNKMDTLEPFGDEFSGMARCPYDAKANALF 177
 119 HNFIVKLLKNDDALFVCGTNAFNSCNKMDTLEPFGDEFSGMARCPYDAKANALF 178
 QY 178 ADGKLYSATVDFLAIDAVITRSLSGEPTLTAVHNSKWLKPEYVQAVDGYTFEPR 237
 179 ADGKLYSATVDFLAIDAVITRSLSGEPTLTAVHNSKWLKPEYVQAVDGYTFEPR 238
 DB 238 ELAIVEYNTGKQVFPFRAVQVCKNDMGSGQVLEKQMTSFLKARLNCVPGDSHFYFNILQ 297
 239 ELAIVEYNTGKQVFPFRAVQVCKNDMGSGQVLEKQMTSFLKARLNCVPGDSHFYFNILQ 298
 QY 298 AVTDVIRINGRDVLTATSTPYNSIPGSAVCAVMDLTASVETGRFKKQKSDSTTPVP 357
 299 AVTDVIRINGRDVLTATSTPYNSIPGSAVCAVMDLTASVETGRFKKQKSDSTTPVP 358
 DB 358 DERVKKPPPGCCAGSSSLERVATSNPPDNLNFKTHPLMDAEPVIFNRPMPJRTWVR 417
 359 DERVKKPPPGCCAGSSSLERVATSNPPDNLNFKTHPLMDAEPVIFNRPMPJRTWVR 418
 QY 418 YELTKIAVDTAGPYQNTTVVFLGSEKGIILFKLARIGNSGF-LNDSLFLEMSVYNSEK 476
 419 YELTKIAVDTAGPYQNTTVVFLGSEKGIILFKLARIGNSGF-LNDSLFLEMSVYNSEK 477
 DB 477 CSYDGEDKRLINGMQLDRASSSLYAFSTCVIKYPLRCERHGGCKTCTASRDPYCGMI 536
 478 CSYDGEDKRLINGMQLDRASSSLYAFSTCVIKYPLRCERHGGCKTCTASRDPYCGMI 537
 QY 537 XEGGACSHLSPPSRILT-FEODIERGNTGLGDCNHSFYALNGHSSLLPSTTSDTAQ 595
 538 XEGGACSHLSPPSRILT-FEODIERGNTGLGDCNHSFYALNGHSSLLPSTTSDTAQ 596
 DB 537 SQ-GSGGRVTPGMLABGYQDTEFGTAHLGCH-----ELDTSTTPD----- 579
 538 SQ-GSGGRVTPGMLABGYQDTEFGTAHLGCH-----ELDTSTTPD----- 580
 QY 580 GYSENRG-----MLD-WKHLL-----DSP-----DSTDP 618
 581 GYSENRG-----MLD-WKHLL-----DSP-----DSTDP 619
 DB 580 -YKIRGPTSDMEVSSSVTMAISPEITPKYIDTWPRKLTSSRKFWVDDEFTSDFTDP 638
 581 -YKIRGPTSDMEVSSSVTMAISPEITPKYIDTWPRKLTSSRKFWVDDEFTSDFTDP 639
 QY 619 LGAVSSHNDKKGYIRRESYLKGHDQYLVPTTLTAIAVILAFYMGAVFSGITVYCVCD -HR 677


```

Db      639  LSGT-----PKGVMEVQSGESNOVMNNVILITVFAAFVIGATAGAAVYCTPDMV 691
Qy      678  RKDVAVQKKEKELTSHRSRSGSMSSVTKLSGLFG---DTQSKDPKPEALITPLMNGKLA 733
Db      692  RKNRKI--HKDAASAGSCDSSGSPAKLNGLPFSPEVKEVQONIDSPKLYSNLTSKELP 749
Qy      734  TPENITAKMLIKADQHLDLTALPTPESTPTLQOKRPSRSGREMEKNOLINACTDMP 793
Db      750  PNDITKSMVNDHKGQPELALPTPESTPTLHOKTLQAMKSHSEKAGH--GASRKETQ 807
Qy      794  MGSPTVPTDLPKASPSHPSVAVVLPITQOQYHEV-----VDOP---KMS 836
Db      808  PFPSSPEPSPFL--SHGHLESAIVLPNATHTDINTSFSNSNAKAEKKLQNIHPLTKSS 865
Qy      837  EVAQMALEDQAATLEKTKIKELHSSKSPN-----HGVNLVENLDTL--PPXP 882
Db      866  KRDRRSVDSRNTL--NDLLKHLNDPNSNPKAIMGDIQMAHQMLMDPMGMEVEPPKVP 923
Qy      883  QREASLGPFGASLSQGLSKRLSMHSSSY--GVYKSYPTNSLTRSHQATTLKRNNT 939
Db      924  NREASLYSPSTLPRNSPTKRVDPPTTPGVMTSLERQRYHNSQQR--HSISAMPK--ML 981
Qy      940  NSNSNSHLSENQSGFGSD--NPPAPQVDSIQVHSSQSPGQAVTAROPSLMAYNSLT-- 996
Db      982  NSPVGVLISQPSMNRGVMPTPTGAKVDYIQ-----GTPVSYHLQPSLSQSSYTSN 1034
Qy      997  ---RSGLKRTPSLKPDVPPKPSFAPLSTSMKP 1025
Db      1035  GTLPRTGLKRTPSLKPDVPPKPSFVPQTPSVRP 1067

```

RESULT 12

```

ID      Q8NFY5      PRELIMINARY;      PRT;      1017 AA.
AC      Q8NFY5;
DT      01-OCT-2002 (TTEMBLrel. 22, Created)
DT      01-OCT-2002 (TTEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TTEMBLrel. 23, Last annotation update)
DE      Semaphorin 6D isoform 3.
GN      SEMA6D.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI TaxID=3606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RA      Ou X., Wei H., Zhai Y., Yu Y., Tang F., He F.;
RT      "Homo sapiens semaphorin 6D isoform 3 (SEMA6D.3) mRNA, complete cds.";
RL      Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF389428; AAM69451.1; -
DR      InterPro; IPR001627; Sema.
DR      Pfam; PFO1403; Sema; 1.
DR      SMART; SM00630; Sema; 1.
SQ      SEQUENCE 1017 AA; 113736 MW; 4D639CEBAD9F2A0 CRC64;

```

Query Match 43.1%; Score 2349; DB 4; Length 1017;
 Best Local Similarity 46.7%; Pred. No. 4.9e-182;
 Matches 503; Conservative 155; Mismatches 302; Indels 116; Gaps 27;

```

Qy      1  MRSALLHYFTLL--HFAAGCFPEDESEPISSHGNTTKOYPVFGKPGKNTTORRLD 57
Db      1  MRVLLCAVITLLMWSQRAVSPEDDEPLNTUVYHSRQYVPRG--RPSNSQ--RLD 58
Qy      58  IQMTIMNGTLIARPHIYVDIDTSHTEBIYCSKULYKWSQADVDTCMKAKHDEC 117
Db      59  FQMLKLRDILYINGRQGVYVNLNEMPKTEVIRNKKULYKWSRQODRENCAMKKGHDEC 118
Qy      118  HNFILVLLKDDALLFVCGTNAFNPSCNRYKMDLLEPGDFSSMACCPYAKKANALF 177
Db      119  HNFILVFPVNDENVFVCGTNAFNPMPCRYLSTLEYDGEBSIGLARCPDPAQRTNALF 178

```

```

Qy      178  ADGKLYSATVVDFLADAVIYRSLSGESPLTRTKVADSXMLEKPEYVQAVDYGYTFEER 237
Db      179  ADGKLYSATVADFLASDAVITRSMGDSALRTIKDSKWIKEPHFLALIEGNYFFER 238
Qy      238  ELAVENMTGKVYFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCSPGDSHFYFNIIQ 297
Db      239  ELAVENHNLGKAVYSRVARIKNDMGSGQVLEKQWTSFLKARLNCSPGDSHFYFDVQ 298
Qy      298  AVTDVIRINGRVVATSTPYNSTPGSAVCAVMDLIDASVFTGRFKEGKSDSTMTVP 357
Db      299  SITDIIQINGIPVYVFTQNLNIPGSAVCAVMDLIEKVKGRKEKTPDSVTAVP 358
Qy      358  DERVEKPRPGCCAGSSSLERVATSNPPDDTNFKTHPIIMDEAVSIPNRMFLATMYR 417
Db      359  EDKVEKPRPGCCAGKGLAAYTSLIDPEPITSLFKSHMLMSAVPPIADEPMTKTRR 418
Qy      418  YRLTIAVDTAAPYQNTVFLGSEKGIILKFLARIGNSG--LNDLCLFLEMVYNSK 476
Db      419  YRLTIAVSDHSAQPYQNTVIFVGSBAGWLVKVLAK--TSPSLNDLVLEIEAYNHAK 476
Qy      477  GSYDVEKRIKMGQDLRASSSLVYAFSTCVIKVLGRGERGKCKTQIARDPYCGMI 536
Db      477  CSAENEEDKVIISLDKDHALVYAFSSCIIRIPLSRCEKRGSCCKSLASRDPYCGM 536
Qy      537  KEGGACSHLSPNSRLT--PEODIERGNTDGLGDCHNSFVALNGHSSSLPSTTSSTAQ 595
Db      537  SQ--GSGCRVTPGMLABGYQDTEFGNTALGCH-----ELPSTSTPD----- 579
Qy      596  GYESRQKMLDKHLNDSPTDPLGAVSHNHQDKGVRESYLKHDOLVPTLLAIYV 655
Db      580  -YKIFGG-----PTS-----GWMEVQSGESNOVMNNVILITCV 612
Qy      656  ILAFTWGAFAFGITVYCVCD--HRKDVAVQKKEKELTSHRSRSGSMSSVTKLSGLFG--- 710
Db      613  FAFVIGATAGTAAVAYCYRDMFVRXNRKI--HKDAASAGSCDSSGSPAKLNGLPFS 670
Qy      711  DTQSKDPKPEALITPLMNGKLAATPGNTAKMLIKADQHLDLTALPTPESTPTLQOKRP 770
Db      671  EYQONIDSPKLYSNLTSKELP--PNNDTCKSMVNDHKGQPELALPTPESTPTLHOKTLQ 730
Qy      771  SRGSRWERNQNLINACTKDMPPMGSPVIFTLPLRASSHIPSVAVVLPITQOQYHEV-- 829
Db      731  AMKSHSEKAGH--GASRKETPQFPSPSPSPFL--SHGHLESAIVLPNATHTDINTSFS 786
Qy      830  -----VDOP---KMEVQMALEDQAATLEKTKIKELHSSKSPN----- 865
Db      787  NSNAHAKAEKKLQNIHPLTKSSSKRDRRSVDSRNTL--NDLLKHLNDPNSNPKAIMGDI 844
Qy      866  ---HGVNLVENLDTL--PPXPQREASLGPFGASLSQGLSKRLSMHSSSY--GVY 916
Db      845  QMAHQMLMDPMGMEVEPPKVPNREASLYSPSTLPRNSPTKRVDPPTTPGVMTSLER 904
Qy      917  KRYPFNSLTRSHQATTLKRNNTNSNSHLSNQGFGSD--NPPAPQVDSIQVHSSQ 975
Db      905  QRGYHNSQQR--HSISAMPK--NLNSPVGVLISQPSMNRGVMPTPTGAKVDYIQ----- 957
Qy      976  PSGQAVTAROPSLMAYNSLT-----RSGLKRTPSLKPDVPPKPSFAPLSTSMKP 1025
Db      958  ---GTPVSYHLQPSLSQSSYTSNGTLPRGLKRTPSLKPDVPPKPSFVPQTPSVRP 1011

```

RESULT 13

```

ID      Q8NFY6      PRELIMINARY;      PRT;      998 AA.
AC      Q8NFY6;
DT      01-OCT-2002 (TTEMBLrel. 22, Created)
DT      01-OCT-2002 (TTEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TTEMBLrel. 23, Last annotation update)
DE      Semaphorin 6D isoform 2.
GN      SEMA6D.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RT Qu X., Zhai Y., Yu Y., Tang F., He F.,
 RA "Homo sapiens semaphorin 6D isoform 2 (SEMA6D.2) mRNA, complete cds."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF389427; AAM69450.1;
 DR InterPro: IPR001627; SEMA.
 DR Pfam: PF01403; SEMA; 1.
 DR SMART, SM00630; SEMA; 1.
 SQ SEQUENCE 998 AA; 111730 MW; 3F46D6872ED5344 CRC64;

Query Match 42.5%; Score 2315.5; DB 4; Length 998;
 Best Local Similarity 45.8%; Pred. No. 2.6e-179;
 Matches 493; Conservative 155; Mismatches 293; Indels 135; Gaps 26;

1 MRSALLLYFTLL--HFAGAFPPEDSEPISSHGNYTKQYVFVGHKPGRTTQRRLD 57
 1 MRVFLCAVILLMTSQLRAVSFPEDDEPLNTVDYHSQYVFRG-RPSGNSQ-HRLD 58
 58 IQMIMNGTLYIARDHIYVDIDTSHTTEIYCSKLTWKSROADVTCRMKGKXDEC 117
 59 FQIMKIRDTLYIAGRDQYVTNLNEMPKTEVI PNKKLTWRSRQODRENCAMKGRKDEC 118
 118 HNEIKVLLKNDALFVCGTNAFNPSCRNYKMDLLEPGDEFSGMARCYDAKHANVAF 177
 119 HNFIFVPRNDENMFVCGTNAFNEMCRYYRSTLEYDSEISGLARCFDARQTNVAF 178
 178 ADGKLYSATVDFLAIDAVIYRSLGESPTLRIVKHSKMLKEPYQAVDYGDIYFFPR 237
 179 ADGKLYSATVDFLAIDAVIYRSLGESPTLRIVKHSKMLKEPYQAVDYGDIYFFPR 238
 238 EIAVEYNTMGKVPFPRVAOVCKNDGSGRVLEKQMTSLKARLNSVGDSEFFENTIQ 297
 239 EIAVEHNNIGKAVYSRVARI CKNDGSGRVLEKHTSFLKARLNSVGDSEFFENTIQ 298
 298 AATDVIRINGRDVLTAFSTPNSIGSAVCAYMDMDIASVFTGRFEQKSPDSTWTPV 357
 299 STIDIIQINGIPVTVGVFTTQLNSIPGSAVCAFSMDIEKVGKGRKEKTPDSVWTAVP 358
 358 DERVPRPGCCAGSSSLERVATSNFPDDTINFIKTHPLMDBAVSIINRPFLEMTWR 417
 359 EDKVPKPRPGCCAKHGLAAYKTSIDFPDETISFIKSHPLMSAVPIADEPWFKTWR 418
 418 YRLTKIADPTAGPYQNHVTVPLSGEKIILIKFLAIGSGF-LNDSLLEEMSVANSK 476
 419 YRLTALSVDSAPYQNYTVIIVGSEAGVLAFLAK--TSPSLNDSVLEETIENYNAK 476
 477 CSYDGEDKRIIMQMDRASSSLYVAFSTCVIKVPIGRCERHSGCKCTCIASRDPYCGMI 536
 477 CSANEDKDVISTLDKDHAIYVAFSSCIIRIPISRCERYSCCKSCIASRDPYCGML 536
 537 KBGACSHLSNPSRLT-FEODIERGNTDGLGDCHNSFVALNGHSSLLPSTTTSDBTAOE 595
 537 SQ-GSGCRVTPGMLAGYEGDETFEGNTAHLGDCHG----- 570
 596 GYSSRGMLDMKLLSPDSTPLGAVSSHNHODKGVIRESLKXHDQVLPVTLTAIAY 655
 571 -----YRWE--VSGES-----NONVHNANVLTICV 593
 656 ILAFVGAAPSGITVYCVCD-HRRKDVAVYORKEKELTHSRGSMGSVTKLSGLFG---- 710
 594 FAAPVLCGAFAGVAVICYRMFPRKRI--HDAISAQSCCTDSSSPALNGLFDSPVYK 651
 711 DTGSKDPKPEPAILTPMENGKLTAPGNTAAMLIKADQHDLDLALPPESTPILQCKRP 770
 652 EYQONIDSPYLNLLTSRKELPPNGDTKSMVMDHSGQPELALPPESTPVLHOKTLQ 711
 771 SRSRERERQNLINACTKMPMGSVIPITDPLASPSHIPSVVLPITQOQYQHEX- 829
 712 AMKSHSEKAGH--GASRKETPOFPBSSPPHSPF--SHGHSATVLPNATHTDINTSFS 767

830 -----VDQF---KMSVQAWELEDQAATLEYKTIKEHLSSKSN----- 865
 768 NSNAHKAKEKLLQNDIHLPTKSSSKRDRRVSVDSENLT--NDLLKHLNPNPNKALMGDI 825
 866 ---HGVNLVENDL---PRVFPQREASLSPGASLSQGTGSKRLNEMHSSSY---GVGY 916
 826 QMAQNLMLDWMGSMSEVPFVPRREASLYSPSTLPNPSYKVDVLTGVPMTSIR 885
 917 KRSYPTNSLTRSHQATTLKRNNTNSNSHLSRNSFGQGD-NPPAPQRYDSIQVHSSQ 975
 886 QRGYHKNSSQR-HSISAMPK-NLNSPNCVLLSRQSPNMRGGMPTPTGAKVDYD----- 938
 976 PSQAAYTVSRQPSLNAYNSLT-----RSGIKRTPSLKPDVPPKPSFAPLSTMKP 1025
 939 --GTPVSVHLQPSLSRQSSYTSNGTLPRTGKTRPSTLKPVDPPKPSFVQTPSVAP 992

RESULT 14
 Q8NFY3 PRELIMINARY; PRT; 1011 AA.

Q8NFY3
 AC Q8NFY3;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Semaphorin 6D isoform 1.
 GN SEMA6D.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RT Qu X., Zhai Y., Yu Y., Tang F., He F.,
 RL "Homo sapiens semaphorin 6D isoform 1 (SEMA6D.1) mRNA, complete cds."
 DR EMBL: AF389430; AAM69453.1;
 DR EMBL: AF389430; AAM69453.1;
 DR InterPro: IPR001627; SEMA.
 DR Pfam: PF01403; SEMA; 1.
 DR SMART, SM00630; SEMA; 1.
 SQ SEQUENCE 1011 AA; 113289 MW; 9D6B8B3633941B89 CRC64;

Query Match 42.5%; Score 2314; DB 4; Length 1011;
 Best Local Similarity 45.4%; Pred. No. 3.5e-179;
 Matches 494; Conservative 155; Mismatches 292; Indels 148; Gaps 26;

1 MRSALLLYFTLL--HFAGAFPPEDSEPISSHGNYTKQYVFVGHKPGRTTQRRLD 57
 1 MRVFLCAVILLMTSQLRAVSFPEDDEPLNTVDYHSQYVFRG-RPSGNSQ-HRLD 58
 58 IQMIMNGTLYIARDHIYVDIDTSHTTEIYCSKLTWKSROADVTCRMKGKXDEC 117
 59 FQIMKIRDTLYIAGRDQYVTNLNEMPKTEVI PNKKLTWRSRQODRENCAMKGRKDEC 118
 118 HNEIKVLLKNDALFVCGTNAFNPSCRNYKMDLLEPGDEFSGMARCYDAKHANVAF 177
 119 HNFIFVPRNDENMFVCGTNAFNEMCRYYRSTLEYDSEISGLARCFDARQTNVAF 178
 178 ADGKLYSATVDFLAIDAVIYRSLGESPTLRIVKHSKMLKEPYQAVDYGDIYFFPR 237
 179 ADGKLYSATVDFLAIDAVIYRSLGESPTLRIVKHSKMLKEPYQAVDYGDIYFFPR 238
 238 EIAVEYNTMGKVPFPRVAOVCKNDGSGRVLEKQMTSLKARLNSVGDSEFFENTIQ 297
 239 EIAVEHNNIGKAVYSRVARI CKNDGSGRVLEKHTSFLKARLNSVGDSEFFENTIQ 298
 298 AATDVIRINGRDVLTAFSTPNSIGSAVCAYMDMDIASVFTGRFEQKSPDSTWTPV 357
 299 STIDIIQINGIPVTVGVFTTQLNSIPGSAVCAFSMDIEKVGKGRKEKTPDSVWTAVP 358
 358 DERVPRPGCCAGSSSLERVATSNFPDDTINFIKTHPLMDBAVSIINRPFLEMTWR 417
 359 EDKVPKPRPGCCAKHGLAAYKTSIDFPDETISFIKSHPLMSAVPIADEPWFKTWR 418

QY	410	YLLTKIAVDTAGLAPGVNHTVFLSEKKIIILKFLAIRGNQSF--LNDLSLEEMSVNSKEK	478
QY	411	YLLTKIAVDTAGLAPGVNHTVFLSEKKIIILKFLAIRGNQSF--LNDLSLEEMSVNSKEK	478
Db	419	YRLTALSVSHSGPQNTVIVFVSGEAMVTLVAK--TSFESLNDVLLBEITATNHA	476
QY	477	CSYDGVEDRRIMGMDLRASSSLVYAFSTCVIKVPLGRCERHGCKATCIASRDPCGM	536
Db	477	CSAENEDKKVILSLDDKHNLVYAFSCIRIPLSCERHYGCKKSCIASDPCGM	536
QY	537	KKGACSHLSPNSRLT-----FEQIERGNTDGLGDGNSFVALANGHSSSL	582
Db	537	SG-GSGCGRTVPMILLTTEFFAFHNHSAEGEODETEFGNTAHLDDCHG-----	583
QY	583	LPSTTSDSGTAGEYSGMGMDMWHLLDSDPDSTGDPAGVSHNHQKKGVIRESYLKG	642
Db	584	-----VRWE--VQSGES-----	593
QY	643	DQLVPTLLALAVILAFVGAVFSGITVYCVCD--HRRKDAVVGQRKEKELTHSRGMS	701
Db	594	NQVHVNNVILITCVPAFVFLGAFIAGVAVCYRDMFVRKRFKI--HKDAESQSCDSG	651
QY	702	VYKLSLFG----DTQSKDPKEALLTELHNGKLTATGNTAKKLIAADQHLDITLPT	757
Db	652	FAKLNGLFDSPVKEYQONIDSEKLYSNLTSRKELPENGDKTSVMMDHRCQPELALPT	711
QY	758	PESTPTLQKKRKPSSRSEMERENQVLINACTKDMPEMSPIPTDLPLRASPSPISVVV	817
Db	712	PESTPTLHOKTLOAMKSHSEKAGH--GASKRETQFPSSPPPHSLP--SHGILPAIV	767
QY	818	LPITQGGYHEX-----VDQP--KVSEVAQNALDQATLEFKTIKHS	860
Db	768	LPNATHDYMTSSNSNAHKAKEKKLQNIHPLTKSSSKDRHRSVDSNTL--NDLLKHLN	825
QY	861	SKSPN-----HGNLVENDSL---PPKPCQEBALGPPGASLSGTGSKLDM	906
Db	826	DFPSNKRALMGDIQNAHQNLMDPMGMSSEVPKVPKNEASLYSPSTLPNNSPTKVDV	885
QY	907	HHSSSY---GVQYKREYPTNSLTSRHOATTLKRNNTSSNSHLSRNOSFRGCD--NP	962
Db	886	PTTPGYPMPSLSEQRGYHNSQSR--HLSAMRK--NLNPNQVLSRQSNMGRGMYTPT	943
QY	963	POEVDSIQVHSSQPSQAVTVSRQPSLNAKNSLT-----RSGLKRTPSLKPDPVPPKSF	101
Db	944	GAKVDTIQ-----GTPSVYHLOSLSRQSYTNGTLPRTGLKRTPSLKPDPVPPKSF	996
QY	1017	APLSTSKMP 1025	
Db	997	VPQTPSVRP 1005	
RESULT 15			
Q9P249			
ID	Q9P249	PRELIMINARY; PRT; 1022 AA.	
AC	Q9P249		
DT	01-OCT-2000	(TrEMBLrel. 15, Created)	
DT	01-OCT-2001	(TrEMBLrel. 18, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Hypothetical protein KIAA1479 (fragment).		
GN	KIAA1479.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20277482; PubMed=10819311;		
RA	Nagase T., Kikuno R., Ishikawa K., Hirotsawa M., Ohara O.;		
RT	"Prediction of the coding sequences of unidentified human		
RT	genes XVII. The complete sequences of 100 new cDNA clones from brain		
RL	DNA Res. 7:143-150(2000)."		
DR	EMBL; AB040912; BAA96003.2; -;		
DR	InterPro; IPR000659; Plexin-like		

DR	InterPro: IPR001627; Sema.
DR	Pfam; PF01403; Sema: 1.
DR	SMART; SM00423; PSI: 1.
DR	SMART; SM00630; Sema: 1.
KW	Hypothetical protein.
FT	NON TER 1
SEQUENCE	1022 AA; 114372 MW; BEAFBD5EA02C69C4 CRC64;
Query Match	42.5%; Score 2314; DB 4; Length 1022;
Best Local Similarity	45.4%; Pred. No. 3.Se-179;
Matches	494; Conservative 155; Mismatches 292; Indels 148; Gaps 26;
QY	1 MRSEALLLPFL---HFAAGPDEDEPISISHGNTTKQYVYVGHKPGSNTTORRLD 57
DB	12 MRVLLCAVILLVWSQLRAPVPEDEDEPLNTVDYHISROYPRVG--RPSNEG--HRLD 69
QY	58 IQMIMNGTLIARDDHIYTDIDTSHTEIYCSKLLTWKSRQADVDTCRMKHKDEC 117
DB	70 FQWLKLRDLYIAGRDQYVTLNINEMPKTEVYIPNKKLTWRSRQDRENCAMKHKDEC 129
QY	118 HNFILVLLKDDALFVCGITAFNPSCRNKMDLLEPGDEFGMAKCPYDAKIANALF 177
DB	130 HNFIKVFPNRDENVEFVCGITAFNPMCRYYRLSTLEVDGEISGLARCPDYARQTNALF 189
QY	178 ADGKLASATYDFALDAVYIRSGESPTRTVGHGSKMLKEPFVQAVDYGDIYTFPR 237
DB	190 ADGKLSTYVADFLASDVIYIRSGDGSALRTIKYDSKIMEPFALAEIGNVYTFPR 249
QY	238 EIAVEYNTMGKVEPRPVAQVCKNDGSGQVLEKQMTSFLKARLNCVPGDSHFYENILQ 297
DB	250 EIAVEHNHLGKAVVSRVARICKNDGSGQVLEKQMTSFLKARLNCVPGDSHFYENILQ 309
QY	298 AVTDVIRINGDVLATSTIYNSIPGSAYAYMDLIASFYTGFRFGKSPDSYTWTFVP 357
DB	310 SITDIQINGIPYVGVFTQINSIPGSAYCAFSGNDIEKVKGRFKQKIPDSYTWTFVP 369
QY	358 DERVKPRPGCCAGASSLERATATNEPDDTLNFKTHPLMDAEVPS-FENPWFRTWVR 417
DB	370 EDKVPKPRPGCCAKHGLAEAKTSDIDFEDTLSTFKSHPLMDSAVPPLADEPWFRTKVR 429
QY	418 YRLTIADVTIAGPYQNTTVFVLGSEKILLKFLARIQNSGF-INDSLFLEMSVYNSK 476
DB	430 YRLTIASVDBAGPYQNTTVIIPVGEAAMVLKVLAK--TSPFLINDSVLEIEIYANNAK 487
QY	477 CSYDVEDDKRIWQMDLRASSSLIYASFTCIYKPLGCEHSGKKKTCIASRDYCGMI 536
DB	488 CSAENEDPKKVIISIQLDKDHALLYVAFSSCIIRIPLSCEBYGGSCKKCIASRDYCGWL 547
QY	537 KEGAGCSHLSPKSRLT-----EEDIERNTGTGLGCHNSFVALNGHSSSL 582
DB	548 SQ-GSGGVTPGMILLTDFPAFHNHSAEGYEDTEFQNTAHLGDCHG----- 594
QY	583 LPSTTSSTAQEGYESRGMIDMKHLIDSPDSTPLGAVSSHNDKKGVIRESYLKH 642
DB	595 -----VRME--VQSGSS----- 604
QY	643 DQLVPTLLALAVILAFWGAVFSGITYYCVCD-HRRKDVAVVQKKEKELTSHRGSMSS 701
DB	605 NQMVHNNVLTICVFAAFVLAGFIAVAVYCYADMFVRKKRKRI--HKDESAQSCDSSGS 662
QY	702 VTKLSGLG---DPOSKDPKEALLPTLNMNGKLATGNTAKMLIKADQHHLDLTALPT 757
DB	663 FAKNLGLPSPVKEVQOINDSPKLYSNLTSSKEKLEPENGDTISKVMHDRGQPELALPT 722
QY	758 PPSPTLLOQRKPRSGRSEMERNQMLINACTKMDPMGSPVITDLPRLASPSHISVYV 817
DB	723 PESTVTLHKTITQAMKSHSEKAHGH--GASRKETLQPFPSPPSPHSPPL--SHGHIIPSAIV 778
QY	818 LPITQOQYQHEV-----VQDP--KMSVDAQALBEOAATLEKTIKEHLS 860
DB	779 LPNACHDYVTSFNSNNAHKAKEKLCNIDHPLTKSSSKDRHRSVSRNTL--NDLLKLHN 836
QY	861 SKSPN-----HGNNLVENIDSL--PPKVPQEBASIGPGASISQIGLSKRLKM 906

Thu Oct 23 17:22:37 2003

us-09-856-681-2.rpt

Page 11

```

Db      DPN$PKAIMEDICMAHONMLDPEG$SEVPKXVPPN$EASLYBPSTLPPN$STKXVDY  896
QY      HHS$Y---GVDYK$YPTN$LT$R$QATLTKRNN$TNS$G$H$SRN$S$FGRCD--NEPPA  962
Db      PTTPEVPMPT$LEROGHYK$SSQR-H$ISAMPK-NL$N$PNCVLL$RQ$SNR$GGYMPPT  954
QY      PQRV$IGVH$Q$P$GQAVT$R$P$SLA$V$SLT-----R$G$KRT$SLK$DVP$PK$SF  1016
Db      955 GAKVDYIQ-----GTPV$YHLP$SL$R$Q$SY$N$G$LT$PRT$GLKRT$SLK$DVP$PK$SF  1007
QY      1017 APL$T$MKP 1025
Db      1008 V$Q$T$SV$P 1016

```

Search completed: October 23, 2003, 17:12:38
Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2003, 17:09:07 ; Search time 17 Seconds
(without alignments)
2849.265 Million cell updates/sec

Title: US-09-856-681-2

Perfect score: 5450
Sequence: 1 MRSEALLLYFTLLHFGAGF.....PPKPSFAPLSTSKPNDACT 1030

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5450	100.0	1030	SM6A_HUMAN	Q9H2E6 homo sapien
2	4346.5	79.8	888	SM6A_MOUSE	Q55464 mus musculu
3	2013.5	36.9	888	SM6B_HUMAN	Q9H2E6 homo sapien
4	1997	36.6	887	SM6B_RAT	Q70141 rattus norv
5	1995	36.6	886	SM6B_MOUSE	Q54951 mus musculu
6	1515.5	27.8	930	SM6C_HUMAN	Q9H312 homo sapien
7	1493.5	27.4	931	SM6C_MOUSE	Q9W133 mus musculu
8	1491	27.4	960	SM6C_RAT	Q9W133 rattus norv
9	971.5	17.8	730	SM1A_SCHAM	Q26473 schistocerc
10	932.5	17.1	771	SM1A_DROME	Q24322 drosophila
11	858	15.7	772	SM3A_RAT	Q63548 rattus norv
12	853	15.7	772	SM3A_MOUSE	Q08605 mus musculu
13	852.5	15.6	771	SM3A_HUMAN	Q14563 homo sapien
14	845	15.5	772	SM3A_MOUSE	Q26972 tribolium c
15	845	15.5	778	SM3A_MOUSE	Q9W133 brachydanio
16	845	15.5	778	SM3A_MOUSE	Q9W133 brachydanio
17	839.5	15.4	772	SM3A_MOUSE	Q9W133 brachydanio
18	830	15.2	1074	SM3A_MOUSE	Q9W133 brachydanio
19	826	15.2	1074	SM3A_MOUSE	Q9W133 brachydanio
20	825.5	15.1	1077	SM3A_MOUSE	Q9W133 brachydanio
21	803	14.7	778	SM3A_MOUSE	Q9W133 brachydanio
22	803	14.7	778	SM3A_MOUSE	Q9W133 brachydanio
23	797	14.6	777	SM3A_MOUSE	Q9W133 brachydanio
24	793.5	14.5	771	SM3A_MOUSE	Q9W133 brachydanio
25	789	14.5	772	SM3A_MOUSE	Q9W133 brachydanio
26	766	14.1	1093	SM3A_MOUSE	Q9W133 brachydanio
27	762.5	14.0	706	SM3A_MOUSE	Q9W133 brachydanio
28	762.5	14.0	706	SM3A_MOUSE	Q9W133 brachydanio
29	752	13.8	751	SM3C_MOUSE	Q9W133 brachydanio
30	751	13.8	861	SM3C_MOUSE	Q9W133 brachydanio
31	750.5	13.8	861	SM3C_MOUSE	Q9W133 brachydanio
32	749	13.7	861	SM3C_MOUSE	Q9W133 brachydanio
33	744.5	13.7	861	SM3C_MOUSE	Q9W133 brachydanio

34	739	13.6	775	SM3B_MOUSE	P70275 mus musculu
35	737	13.5	751	SM3C_MOUSE	Q62181 mus musculu
36	735.5	13.5	862	SM4D_HUMAN	Q92854 homo sapien
37	733.5	13.5	785	SM3C_CHICK	Q42237 gallus gall
38	733	13.4	751	SM3C_HUMAN	Q99985 homo sapien
39	700.5	12.9	766	SM27_BRARE	Q9YH44 brachydanio
40	692	12.7	834	SM4C_MOUSE	Q64151 mus musculu
41	681	12.5	832	SM4B_HUMAN	Q9H351 homo sapien
42	671	12.3	761	SM4A_HUMAN	Q9H351 homo sapien
43	664	12.2	776	SM4F_RAT	Q92143 rattus norv
44	658	12.1	777	SM4F_MOUSE	Q92123 mus musculu
45	656	12.0	782	SM4B_MOUSE	Q62179 mus musculu

ALIGNMENTS

RESULT 1
ID SM6A_HUMAN STANDARD; PRT; 1030 AA.
AC Q9H2E6; Q9P2H9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1) (SEMA6A-1).
GN SEMA6A OR KIAA1368.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INTERACTION WITH EVL.
RX MEDLINE=20564339; PubMed=10939894;
RA Klottermann A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated phosphoprotein-like protein (EVL) via a novel carboxyl-terminal tyrosine-like domain.";
RT J. Biol. Chem. 275:39647-39653(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro.";
RT DNA Res. 7:65-73(2000).
CC -!- FUNCTION: Can act as repulsive axon guidance cues. May play a role in channeling sympathetic axons into the sympathetic chains and controlling the temporal sequence of sympathetic target innervation (by similarity).
CC -!- SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9H2E6-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9H2E6-2; Sequence=VSP_007113;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL: AF279656; AAG29378.1; -; ALT_INIT.
 DR EMBL: AB037789; BAA92606.1; SEMA6A.
 DR Genew: H0NC10738; SEMA6A.
 DR MIM: 605885; -;
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam: Pf01403; Sema; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR Signal: Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein; Alternative splicing.
 KW SIGNAL
 FT CHAIN 1 18 POTENTIAL.
 FT DOMAIN 19 1030 SEMAPHORIN 6A.
 FT TRANSLEM 650 670 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 671 1030 POTENTIAL.
 FT DOMAIN 56 491 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 792 819 SEMA.
 FT CARBOHYD 33 33 PRO-RICH.
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 576 576 N -> NDSITPLEDNMSYNTVY (in isoform 2).
 FT /FTID=VSP_007113.
 SC SEQUENCE 1030 AA; 114368 MW; A57B79C10AECAB34 CRC64;

Query Match 100.0%; Score 5450; DB 1; Length 1030;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1030; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSEALLLYFTLLFAAGAFPEDESEPISSHGNTKQYPPVFGHKPRNTTQRRLDIOM 60
 DB 1 MRSEALLLYFTLLFAAGAFPEDESEPISSHGNTKQYPPVFGHKPRNTTQRRLDIOM 60
 QY 61 IMINMGTYTARDDHYTVDIDTSHTBEIYCSKLTWKSROADVDTCKMKKXDECHNF 120
 DB 61 IMINMGTYTARDDHYTVDIDTSHTBEIYCSKLTWKSROADVDTCKMKKXDECHNF 120
 QY 121 IKVLLKKNDLAFVCGTNAEPSCRYNMDLEPFGEFSSMARCPCYDAKANVALPADG 180
 DB 121 IKVLLKKNDLAFVCGTNAEPSCRYNMDLEPFGEFSSMARCPCYDAKANVALPADG 180
 QY 181 KLYSATVDFLAIDAVIYRSLSGESPTLRTVHDSKWLKEPYVOADVGYDYTFEPEIA 240
 DB 181 KLYSATVDFLAIDAVIYRSLSGESPTLRTVHDSKWLKEPYVOADVGYDYTFEPEIA 240
 QY 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCSPGDSHFEYFNLQAVT 300
 DB 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCSPGDSHFEYFNLQAVT 300
 QY 301 DVIRNGRDVLAFTSTPYNSIPGSAVCAYDMLDIASVFTGRFKQKSPDSTWTPVPPER 360
 DB 301 DVIRNGRDVLAFTSTPYNSIPGSAVCAYDMLDIASVFTGRFKQKSPDSTWTPVPPER 360
 QY 361 VPKPRPGCCAGSSSLERYATSNPPDDTLNFKTHPLMDEAVPSIFENPWLRTWRYRL 420
 DB 361 VPKPRPGCCAGSSSLERYATSNPPDDTLNFKTHPLMDEAVPSIFENPWLRTWRYRL 420
 QY 421 TKIADVTAAGPYQNTTVFPLGSEKGIILKFLARINSGLFNDSTLEMSVYNSKSGYD 480
 DB 421 TKIADVTAAGPYQNTTVFPLGSEKGIILKFLARINSGLFNDSTLEMSVYNSKSGYD 480
 QY 481 GVEDRIMGMDLRASSSLIYVAFSTCVIKVPLGRCERHCKCKTKCIASRDPIYCGMIKGG 540
 DB 481 GVEDRIMGMDLRASSSLIYVAFSTCVIKVPLGRCERHCKCKTKCIASRDPIYCGMIKGG 540
 QY 541 ACSHSPNSRLTFEODIERGNTDGLGDCNNSVVALNGHSSLLPSTTSDSTAGQYSR 600
 DB 541 ACSHSPNSRLTFEODIERGNTDGLGDCNNSVVALNGHSSLLPSTTSDSTAGQYSR 600
 QY 601 GGMIDMKHLIDSPDSTDPDGLGAVSNHNDKQKVIKGDQVPTVLLAIVLAFV 660
 DB 601 GGMIDMKHLIDSPDSTDPDGLGAVSNHNDKQKVIKGDQVPTVLLAIVLAFV 660

DB 601 GGMIDMKHLIDSPDSTDPDGLGAVSNHNDKQKVIKGDQVPTVLLAIVLAFV 660
 QY 661 MGAVESGITVYCVCDHRRDVAVVOREKELTHSRGSSVYTKISGLFGDTQSDPKRE 720
 DB 661 MGAVESGITVYCVCDHRRDVAVVOREKELTHSRGSSVYTKISGLFGDTQSDPKRE 720
 QY 721 ALLTPLENGKATAGNTAKXMKIKADQHLDTLALPTPESTPTLOQKRPSGRSEMER 780
 DB 721 ALLTPLENGKATAGNTAKXMKIKADQHLDTLALPTPESTPTLOQKRPSGRSEMER 780
 QY 781 QNLINACTKMPMPSPVPTDPLRASPSPHPSVVLPTTQGYCHEYVDQKXSEVAQ 840
 DB 781 QNLINACTKMPMPSPVPTDPLRASPSPHPSVVLPTTQGYCHEYVDQKXSEVAQ 840
 QY 841 MALEDQATLEYKTKIHELSSKSPHGVNLVENDLSLPKVPVOREASLGPAGSLISQTL 900
 DB 841 MALEDQATLEYKTKIHELSSKSPHGVNLVENDLSLPKVPVOREASLGPAGSLISQTL 900
 QY 901 SKRLMHSSSYGVYKRSYPTNSLTRSHOATTKNNNSNSSHLSRNSQFGRDNP 960
 DB 901 SKRLMHSSSYGVYKRSYPTNSLTRSHOATTKNNNSNSSHLSRNSQFGRDNP 960
 QY 961 PAPQVDSIOVHSSQPSGOAVTVSRQPSINAYNSLTRSGLKPTPSLKPDVPPSPAPLS 1020
 DB 961 PAPQVDSIOVHSSQPSGOAVTVSRQPSINAYNSLTRSGLKPTPSLKPDVPPSPAPLS 1020
 QY 1021 TSMKPNDACT 1030
 DB 1021 TSMKPNDACT 1030

RESULT 2
 SM6A MOUSE
 ID SM6A MOUSE STANDARD; PRT; 888 AA.
 AC 035464;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1)
 GN SEMA6A OR SEMAQ.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=97348468; Pubmed=9204478.
 RA Zhou L., White F.A., Lentz S.I., Wright D.E., Fisher D.A.,
 RT "Cloning and expression of a novel murine semaphorin with structural
 RT similarity to insect semaphorin I".
 RL Mol. Cell. Neurosci. 9:26-41(1997).
 RN [2]
 RP INTERACTION WITH EVL.
 RX MEDLINE=20564339; Pubmed=10993894;
 RA Klosternann A., Lutz B., Gertler F., Behl C.;
 RT "The orthologous human and murine semaphorin 6A-1 proteins
 RT (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated
 RT phosphoprotein-1-like protein (EVL) via a novel carboxyl-terminal
 RT zyxin-like domain".
 RL J. Biol. Chem. 275:39647-39653(2000).
 CC -1- FUNCTION: Can act as repulsive axon guidance cues. May play a role
 CC in channeling sympathetic axons into the sympathetic chains and
 CC controlling the temporal sequence of sympathetic target
 CC innervation.
 CC -1- SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: PARTICULARLY HIGH LEVELS IN SPINAL CORD,
 CC CEREBELLUM, MENINGEAL PHYLLOID AND INFERIOR COLICULUS,
 CC DENCEPHALON, OLFACTORY BULB, AND EYE.
 CC -1- DEVELOPMENTAL STAGE: TEMPORALLY AND SPATIALLY REGULATED DURING

DEVELOPMENT
 -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 -1- SIMILARITY: Contains 1 Sema domain.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF030430; AAB86408.1; -
 MGI; MGI:1203727; Sema6a.
 InterPro; IPR003659; Plexin-like.
 InterPro; IPR01627; Sema.
 Pfam; PF01403; Sema; 1.
 SMART; SM00423; PST; 1.
 SMART; SM00630; Sema; 1.
 Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 Developmental protein.
 KMW
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 888 SEMAPHORIN 6A.
 FT DOMAIN 19 649 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 650 670 POTENTIAL.
 FT DOMAIN 671 888 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 56 491 SEMA.
 FT DOMAIN 792 819 PRO-RICH.
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 888 AA; 99075 MW; C7094FC2B345C57 CRC64;

Query Match 79.8%; Score 4346.5; DB 1; Length 888;
 Best local similarity 94.4%; Pred. No. 3.5e-272;
 Matches 816; Conservative 24; Mismatches 23; Indels 1; Gaps 1;

QY 1 MRSEKLLVFTLLHAGAGPEPDSPPISISGNTYKQPVFGHKGKNTQRRHLDIQM 60
 DB 1 MRPALLLCTLLHAGAGPEPDSPPISISGNTYKQPVFGHKGKNTQRRHLDIQM 60
 QY 61 IMINAGTYIARADITVDIDTSHTEELYSKLTWRSRQADVDTCKMGKHKDCENF 120
 DB 61 IMINRRTLYVARDITVDIDTSHTEELYSKLTWRSRQADVDTCKMGKHKDCENF 120
 QY 121 IKVLLKKDDALFVCGTNAFNSCGNYKMDTELEPGDESGMARCPYDAKANVALLFADG 180
 DB 121 IKVLLKKDDTLFVCGTNAFNSCGNYRVDTLETFGDESGMARCPYDAKANVALLFADG 180
 QY 181 KIYSAITVDELAIIDAVIYRSIGESPTLRTVHKDSKMLKEPYVQAVDGYIFFFREIA 240
 DB 181 KIYSAITVDELAIIDAVIYRSIGESPTLRTVHKDSKMLKEPYVQAVDGYIFFFREIA 240
 QY 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVT 300
 DB 241 VEYNTMGKVPFPRVAQVCKNDMGSGQVLEKQWTSFLKARLNCVPGDSHFYFNILQAVT 300
 QY 301 DVIRINGRDVLAITSTPYNSIPGSAVCAVMDLIASVTFGRPKGKSDSTWTPPPDR 360
 DB 301 DVIRINGRDVLAITSTPYNSIPGSAVCAVMDLIASVTFGRPKGKSDSTWTPPPDR 360
 QY 361 VKPRPGCGAGSSSEKATSNPEPDTLNFIKTHPLMDCAVPSINRPMFLRTWRYEL 420
 DB 361 VKPRPGCGAGSSSEKATSNPEPDTLNFIKTHPLMDCAVPSINRPMFLRTWRYEL 420
 QY 421 TKIAVDTAAGPYONHTVFLGSEKGIILKFLARIGNSGFLNDSLPLEKSVYNSEKCYSD 480
 DB 421 TKIAVDTAAGPYONHTVFLGSEKGIILKFLARIGNSGFLNDSLPLEKSVYNSEKCYSD 480
 QY 481 GVEDKRIKMGQUDRASGLIYAFSTCVIKVPLGRCRHRGCKKCTCLASDPYCGWKIEEG 540

DB 481 GVEDKRIKMGQUDRASGLIYAFSTCVIKVPLGRCRHRGCKKCTCLASDPYCGWKIEEG 540
 QY 541 ACSHLSPNRRLTPEQDIERKNTDGLDCNSFALNGHSSLLPSTTSDSTOEYSESR 600
 DB 541 ACSHLSPNRRLTPEQDIERKNTDGLDCNSFALNGHSSLLPSTTSDSTOEYSESR 600
 QY 601 GGWLDKHLHLLSDSDTDPGAVSSHHNODKGVYRESYLGKHQDLPVTLAIAVILAFV 660
 DB 601 GGWLDKHLHLLSDSDTDPGAVSSHHNODKGVYRESYLGKHQDLPVTLAIAVILAFV 660
 QY 661 MGAVFSGITVYCYCDHRKQVAVYQREKELTHSRGSMSSVTKLGLFEDTQSKDPKPE 720
 DB 661 MGAVFSGITVYCYCDHRKQVAVYQREKELTHSRGSMSSVTKLGLFEDTQSKDPKPE 720
 QY 721 AITPLMENGKATPGNTAKMLIKAOHDLTALPPESTPTLQOKRKRSPSRWEERN 780
 DB 721 AITPLMENGKATPGNTAKMLIKAOHDLTALPPESTPTLQOKRKRSPSRWEERN 780
 QY 781 ONLINCTKMPMPGSPVPTDILPLRASPSPHPSVVLPTTQOGYQHEVVDQPMSE-VA 839
 DB 781 ONLINCTKMPMPGSPVPTDILPLRASPSPHPSVVLPTTQOGYQHEVVDQPMSE-VA 839
 QY 840 QMALEDQATLEKTKIKELSSKS 863
 DB 841 QMALEDQATLEKTKIKELSSKS 864

RESULT 3
 SM6B_HUMAN STANDARD; PRT; 888 AA.
 ID SM6B_HUMAN
 AC Q9H3T3; Q9NRK9;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 6B precursor (Semaphorin Z) (Sema Z).
 GN SEMA6B OR SEMAZ.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC Tissue=Brain;
 RA Kimura T., Ishida H.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21248680; PubMed=11350127;
 RA Correa R.G., Sasahara R.M., Bengtson M.H., Katayama M.L.H.,
 RA Salim A.C.M., Brentani M.W., Sogayar M.C., de Souza S.J.,
 RA Simpson A.J.G.;
 RT "Human semaphorin 6b";
 RL Genomics 73:343-348(2001).
 CC -1- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS
 CC SYSTEM DEVELOPMENT (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing. Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=Q9H3T3-1; Sequence=Displayed;
 CC Name=2; Synonyms=68.1;
 CC IsoId=Q9H3T3-2; Sequence=VSP_006044, VSP_006045;
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR Signal: Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein.
 FT SIGNAL 1 26
 FT CHAIN 27 887
 FT DOMAIN 27 605
 FT TRANSMEM 606 626
 FT DOMAIN 627 887
 FT DOMAIN 239 549
 FT CARBOHYD 75 75
 FT CARBOHYD 156 156
 FT CARBOHYD 168 168
 FT CARBOHYD 292 292
 FT CARBOHYD 387 387
 FT CARBOHYD 442 442
 FT CARBOHYD 463 463
 FT SEQUENCE 887 AA; 95752 MW; 095433FP202CD301 CRC64;
 Query Match 36.6%; Score 1997; DB 1; Length 887;
 Best Local Similarity 42.7%; Pred. No. 7,1e-121; Indels 190; Gaps 25;
 Matches 433; Conservative 132; Mismatches 260;
 QY 2 RSEALLLEFLLHFGAGFPEDESEFISHGNYTKQYFVGHKGRNTGR--HRLDIQ 59
 DB 10 RPAFLFLLLLRVTHGFFPDEPPLSVAPRDYLSHYVFVCGSGRGLTPAGADLNIQ 69
 QY 60 MIMMNGFLYIARDHIYTDIDTSHTEIYCSKULTKSKQADVDTCRMGKHDECHN 119
 DB 70 RYLRVNRFLTIGDRDNLQVELEPSTSTELRYQRKLTFRSNSSDIDVCRMGKQGECHN 129
 QY 120 FIKYLLKXNDLAFVCGNAPNCRNYKMDLEPFGEFSGMARCPYDAKIANVALPAD 179
 DB 130 FVKVLLDESTLLFVCGSNAPNCRNYKMDLEPFGEFSGMARCPYDAKIANVALPAD 189
 QY 180 GKLVSATVDFLAIDAVIYRSLSGESPTLRVYKSKMLKEFYVQAVDYGYIYFFPREI 239
 DB 190 GMLFATVDFLAIDAVIYRSLSGESPTLRVYKSKMLKEFYVQAVDYGYIYFFPREI 249
 QY 240 AVEYNTMGVYPRPAQVCKNMGSQRLTEKOMTSFLKARLNCVPGSHHYFETILOAV 299
 DB 250 AMEFYLEKVVVSRVAVCKNDVGSPLYLEKQWTSFLKARLNCVPGSHHYFETILOAV 309
 QY 300 TVVIRLNGDVLAFTSTPNSIPGSAVCAYMDLIASVFTGRFKEOKSPDSTWTPVDE 359
 DB 310 TGVVSLGRPVILAVFTSTPNSIPGSAVCAYMDLIASVFTGRFKEOKSPDSTWTPVDE 369
 QY 360 RYVPRPGCCAGSSSLERYATSNFPDDTLNFKTHPLMDEAVPSIFENRPFELRTWYR 419
 DB 370 QVPRPRPGCCAGSSSLERYATSNFPDDTLNFKTHPLMDEAVPSIFENRPFELRTWYR 427
 QY 420 LTKIADVTAAGPYQNTVVFILGSEKIIILKFLAR--IGNSGLNDLFLFLEMSVYNSK 477
 DB 428 LTKIADVTAAGPYQNTVVFILGSEKIIILKFLAR--IGNSGLNDLFLFLEMSVYNSK 487
 QY 478 ---SYGVGDKRMGMQDLBASGLVAVASTCVIKYPLGRCERHAGCKTKTCLASRDPYG 534
 DB 488 GSSSSAGMGQRLSLSELDASGLLAAPRCVYRPPVAVRCQYSGCMNCIGSQDPYOG 547
 QY 535 WIKGGACSHLSFNSRLTFEODIERGNTDGLGDCHNSFYALNGHSSLLPSTTSDSTQA 594
 DB 548 WARD-GSCIFLRPGTSATFEQDVSASTSLGDC----- 580
 QY 595 EGVESRGMLDMKHLSDPSTPLAIVSHNQDKKGVIRESYLKGHDVLEVTLLATA 654
 DB 581 -----TGLRSLSDRACLVSNLLVTS 604
 QY 655 VILAFVWAGVFSGLTV--YVCCHRRKQVAVVORKEKE--LTHSRGSGMSVTXL----- 705
 DB 605 SVAAVAVGAVVSGSVGMVGLRERELA--RKDKVAILAHGSEAVSVSLGRGRT 662

QY 706 -SGLFDTQSKDKPKREALITPLMHNGKLTAPGNTAKMLIKADQHHDLTLPPESTPT 763
 DB 663 GTGGRGAGGPGGPPPALAPLMONG-----TRALLHGGPHDLSGLPTPEQTP- 715
 QY 764 LQGRKRSR-----GSREMERONLINACTK-----DMPGSPVITPD-- 802
 DB 716 LPGRKLPFTPHALAGPRAMDSHLLSASASTLLLAHTRAPEQP-----VPTESG 770
 QY 803 -----LPLRASHIPSVVVLPITQGYGHEVYDQPKMSEVAMALDDQAATLEYTK 856
 DB 771 PESRLCAPRSCRAHPGDPPLTP-----HASPDRRRVSAPTGPLDSSSVG----- 816
 QY 857 EHLSSKSPNNGVNLVENLDELTP--PKYQREASL-----GPPGASLSQGTGLSKRLEMHS 909
 DB 817 -----DLPGWSPATSSLRPPGHPPTALART-----HT 849
 QY 910 SSYGVYKRSYPTNSLTRSHQATTLKRNNTSSNSHLSRNSQSGRD-NEPPAP 963
 DB 850 FNSG-----EAPFGHRRRRA-----PADSTHL--LPGGTGERTAPVP 887
 RESULT 5
 SM6_MOUSE STANDARD; PRT; 886 AA.
 ID SM6_MOUSE
 AC 054951;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 6B precursor (Semaphorin VIB) (Sema VIB) (Semaphorin N) (Sema N).
 GN SEMA6B OR SEMAN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98027184; PubMed=9361278;
 RA Eckhardt F., Behr O., Calutci E., Yonezawa K., Nishimoto I.,
 RA Fishman M.C.;
 RT "A novel transmembrane semaphorin can bind c-src".
 RL Mol. Cell. Neurosci. 9:409-419(1997).
 CC - SUBUNIT: HOMODIMER. BINDS SPECIFICALLY THE SH3 DOMAIN OF THE PROTOONCOGENE C-SRC.
 CC - TISSUE SPECIFICITY: Type I membrane protein.
 CC - SUBCELLULAR LOCATION: DURING DEVELOPMENT IT IS EXPRESSED IN MUSCLE. IN ADULTHOOD, IT IS EXPRESSED UBQUITOUSLY.
 CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC - SIMILARITY: Contains 1 Sema domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF036585; AAC00493.1; -
 CC MGD: MGI:102889; Sema6b.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR Signal: Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 KW Developmental protein.
 FT SIGNAL 1 26
 FT CHAIN 27 886
 FT DOMAIN 27 605
 FT TRANSMEM 606 626
 FT SEQUENCE 886 AA; 95752 MW; 095433FP202CD301 CRC64;
 Query Match 36.6%; Score 1997; DB 1; Length 886;
 Best Local Similarity 42.7%; Pred. No. 7,1e-121; Indels 190; Gaps 25;
 Matches 433; Conservative 132; Mismatches 260;
 QY 2 RSEALLLEFLLHFGAGFPEDESEFISHGNYTKQYFVGHKGRNTGR--HRLDIQ 59
 DB 10 RPAFLFLLLLRVTHGFFPDEPPLSVAPRDYLSHYVFVCGSGRGLTPAGADLNIQ 69
 QY 60 MIMMNGFLYIARDHIYTDIDTSHTEIYCSKULTKSKQADVDTCRMGKHDECHN 119
 DB 70 RYLRVNRFLTIGDRDNLQVELEPSTSTELRYQRKLTFRSNSSDIDVCRMGKQGECHN 129
 QY 120 FIKYLLKXNDLAFVCGNAPNCRNYKMDLEPFGEFSGMARCPYDAKIANVALPAD 179
 DB 130 FVKVLLDESTLLFVCGSNAPNCRNYKMDLEPFGEFSGMARCPYDAKIANVALPAD 189
 QY 180 GKLVSATVDFLAIDAVIYRSLSGESPTLRVYKSKMLKEFYVQAVDYGYIYFFPREI 239
 DB 190 GMLFATVDFLAIDAVIYRSLSGESPTLRVYKSKMLKEFYVQAVDYGYIYFFPREI 249
 QY 240 AVEYNTMGVYPRPAQVCKNMGSQRLTEKOMTSFLKARLNCVPGSHHYFETILOAV 299
 DB 250 AMEFYLEKVVVSRVAVCKNDVGSPLYLEKQWTSFLKARLNCVPGSHHYFETILOAV 309
 QY 300 TVVIRLNGDVLAFTSTPNSIPGSAVCAYMDLIASVFTGRFKEOKSPDSTWTPVDE 359
 DB 310 TGVVSLGRPVILAVFTSTPNSIPGSAVCAYMDLIASVFTGRFKEOKSPDSTWTPVDE 369
 QY 360 RYVPRPGCCAGSSSLERYATSNFPDDTLNFKTHPLMDEAVPSIFENRPFELRTWYR 419
 DB 370 QVPRPRPGCCAGSSSLERYATSNFPDDTLNFKTHPLMDEAVPSIFENRPFELRTWYR 427
 QY 420 LTKIADVTAAGPYQNTVVFILGSEKIIILKFLAR--IGNSGLNDLFLFLEMSVYNSK 477
 DB 428 LTKIADVTAAGPYQNTVVFILGSEKIIILKFLAR--IGNSGLNDLFLFLEMSVYNSK 487
 QY 478 ---SYGVGDKRMGMQDLBASGLVAVASTCVIKYPLGRCERHAGCKTKTCLASRDPYG 534
 DB 488 GSSSSAGMGQRLSLSELDASGLLAAPRCVYRPPVAVRCQYSGCMNCIGSQDPYOG 547
 QY 535 WIKGGACSHLSFNSRLTFEODIERGNTDGLGDCHNSFYALNGHSSLLPSTTSDSTQA 594
 DB 548 WARD-GSCIFLRPGTSATFEQDVSASTSLGDC----- 580
 QY 595 EGVESRGMLDMKHLSDPSTPLAIVSHNQDKKGVIRESYLKGHDVLEVTLLATA 654
 DB 581 -----TGLRSLSDRACLVSNLLVTS 604
 QY 655 VILAFVWAGVFSGLTV--YVCCHRRKQVAVVORKEKE--LTHSRGSGMSVTXL----- 705
 DB 605 SVAAVAVGAVVSGSVGMVGLRERELA--RKDKVAILAHGSEAVSVSLGRGRT 662

FT DOMAIN 627 886 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 239 549 SEMA.
 FT DOMAIN 751 754 POLY-LEU.
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SC SEQUENCE 886 AA; 35466 MW; E5F56D125CDA574D CRC64;

Query Match 36.6%; Score 1995; DB 1; Length 886;
 Best Local Similarity 43.7%; Pred. No. 9.5e-121;
 Matches 438; Conservative 132; Mismatches 257; Indels 176; Gaps 28;

QY 8 LYEFTLLHFAAG--PPEDESEPTISHGNYTKQYFVGHKRGRTTOR--HRLDIOMT 62
 DB 13 LSFLLLLLVGYTGLFPEEPPLSVAPRYLSHYFVSSGGRLAAGAEMLNQRL 72
 QY 63 INNGTLIYIARDHIVYVIDIDTSHTEITYSSKLTWKSQADVDTCMKGKHDECHNFIK 122
 DB 73 RVNRITLFIQBRDMLYQVELEPSTSTELRYQRKLTWNSPNSDIDVCMKKGKECECNFVK 132
 QY 123 VLLKKNDDALFVCGTAAPNPSCKNYKMDLPEFGDESGMARCPYAKHANYALFADGKL 182
 DB 133 VLLRDESTLFVCGSNAFNPICANYSMDTLQLGDSISGMARCPDIPKIANVALFEDGML 192
 QY 183 YSATVDFDLADAVIYRSLGESPTLRTVKHDSKMLKEPFVQADVDYDYIYFFEREIAVE 242
 DB 193 FTLATVDFDLADAVIYRSLGDRPTLRTVKHDSKMLKEPFVQADVDYDYIYFFEREIAVE 252
 QY 243 YNTMGKVPFPPVAVOCVCKNMGSGQRYLEKOWTFLKARLNCSPGSHHYFNILQAVTV 302
 DB 253 FNYLEKVVSRVAVRCKNDVGSPRYLEKOWTFLKARLNCSPGSHHYFNILQAVTV 312
 QY 303 IIRNGEDVIAITSTYNSIPGSAVCAYMDLJIAFTGRFQKQKPSDTPWTPVERBP 362
 DB 313 VSIAGGPRVIAVSTBSNIPGSAVCAYMDLJIAFTGRFQKQKPSDTPWTPVERBP 372
 QY 363 KPRPGCAGSSSLERYATSNFPDDTLNFIKTHPLMDEAVPSIFNBPMLRTVRYRLTK 422
 DB 373 RPRPGCAARGM--QYNASGALDELINFIKTHPLMDEAVPSIGHSPWIRTLMRQHLR 430
 QY 423 IAVDTAAGPYQNTVVFVLESEKGIILKFLAR--IGNSGFINDLFLSEMSVYNSEK--- 477
 DB 431 VAVDVGAGPWGNOTIYFLGSEAGVILFKVPNASVSGTTPSIFLEEFETYPDRGGR 490
 QY 478 SYDGEDEKRIWGQDRASSSLVAFSTCVIKYPLGCEHNGCKKTCIASRDPYCGWIK 537
 DB 491 SSGGEMGQRLSLJELDAASGGLLAAPRCVAVRPVAVRCQLYSGCMKKGCSQDPYCGMAP 550
 QY 538 EGGACSHLSPNSRLTFEODIERGNTDGLGDCHNSFVALNGHSSSLPSTTSDSTAQEGY 597
 DB 551 D-GSCIFLRPGTSAITFEQDVSGASTSLGDC----- 580
 QY 598 ESRGMLDMKHLDPDSTPLGAVSHNODKKGVRESYLGKHQDLYVYVTLTALAIVL 657
 DB 581 -----TGLRAHSLSDRAGLVSVNLVLTSSVA 607
 QY 658 AFVWGAFFSGITV--YVCDDRKKDVAVVORKEKE--LTHSRPGMSVTKLSGLFGDTOS 714
 DB 608 AFVWGAFFSGITV--YVCDDRKKDVAVVORKEKE--LTHSRPGMSVTKLSGLFGDTOS 714
 QY 715 KDP-----KREAILTPIMHNGKLTATPNTAKMLIKADQHLDLTALPTPESTP 762
 DB 662 TGPGGAGGAGGPGPEALALPIMQGM-----TKAALHGGPHDLDTGLLPTPEPTP 715
 QY 763 TLQCKRP-----SGSREMENQULINACTCMMPMSPVPTL-----PLRASP 811
 DB 716 -LPCKRLPFPHPAHALGSRAMHSHALSSAS-----TSLLIAPRASEQ- 762
 QY 812 IPSVAVLPITQGVQHEVDDP--KMSVAVQMLLEQATLEVTKTKEHLSSKSPMHGV 868

DB 763 -PQVPAEPGPE-----SRLCAPRGASHPCDFLTPHASPDRRRVVSAPTGLDPVVG- 815
 QY 869 NLVENLDSLP-PKQVQREASL-----GPPASLSQTLSKRLHSSSVGVYKAGVP 921
 DB 816 -----DLPLPQWSPSPATSLRRPQPGPFAAARRT-----HTFNSG-----EARP 856
 QY 922 TNSLRSHQATTLKRNNTSSNSHLSNQSFGKRD-NPPAP 963
 DB 857 GGHRRFRHP-----PADSTHL--LPGCTERTAPFVP 886

RESULT 6

SMGC_HUMAN STANDARD; PRT; 930 AA.
 ID SMGC_HUMAN Q9H3T2; Q8WX19; Q8WXU0;

AC Q9H3T2; Q8WX19; Q8WXU0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
 GN SEMA6C OR SEMAY.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RA Kimura T., Ishida H.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC TISSUE=Brain;
 RA Ou X., Zhai Y., Wei H., Zhang C., Xing G., Yu Y., Wu S., Zhang Y.,
 RA OuYang S., Zhou G., He F.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE A STOP SIGNAL FOR THE DORSAL ROOT GANGLION
 CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER
 CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF
 CC NEURONAL CONNECTIONS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1; Synonyms=Short 1;
 CC Name=2; Synonyms=Short 2; Sequence=Displayed;
 CC Name=3; Synonyms=Long;
 CC IsoId=Q9H3T2-3; Sequence=VSP_006047;
 CC IsoId=Q9H3T2-2; Sequence=VSP_006046; VSP_006047;
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC
 CC EMBL; AB022434; BAB20670.1; -
 CC EMBL; AF339153; AAL72098.1; -
 CC EMBL; AF339152; AAL72099.1; -
 CC EMBL; AF339154; AAL72100.1; -
 CC Genew; HGNC:10740; SEMA6C.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
 CC developmental protein; Alternative splicing.
 CC SIGNAL 1 24
 FT CHAIN 25 930 SEMAPHORIN 6C.

FT DOMAIN 25 604 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 605 625 POTENTIAL.
 FT DOMAIN 626 930 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 233 540 SEMA.
 FT DOMAIN 662 667 POLY-PRO.
 FT CARBOHYD 752 755 POLY-PRO. (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 184 223 Missing (in isoform 2).
 FT VARSPLIC 586 586 /FtId=VSP 006046.
 FT VARSPLIC 586 586 Y-> YVDFGPGSPGTSPPSDMHPPOSSTLGVNTR
 (in isoform 2 and isoform 3).
 FT CONFLICT 125 125 /FtId=VSP 006047.
 FT CONFLICT 252 252 I-> V (IN REF. 2; AAL72099).
 FT CONFLICT 455 455 R-> K (IN REF. 1).
 FT CONFLICT 455 455 P-> T (IN REF. 1).
 SQ SEQUENCE 930 AA; 96682 MW; 8AF8814ADB84C88 CRC64;

Query Match 27.8%; Score 1515.5; DB 1; Length 930;
 Best Local Similarity 35.8%; Pred. No. 7,7e-90;
 Matches 380; Conservative 136; Mismatches 356; Indels 189; Gaps 32;

QY 6 LLLYFTLLHFAAGAFPEDESEPISSHGNVYKQYVYFVGHKPKGRNTQNRHLDIQMIMIMN 65
 DB 13 LLLLSLSPH-TQAPFPDPLPLISDLQSTPLSWFRGLDADAVAL-GLDFQRFLLIN 70
 QY 66 GTLYIAARDHITYVDIDTSHT-EEIYCSKLTWKSROADVDTCKRKHGKHDECHNFIVL 124
 DB 71 RLLVAARDHVFSPDLQEBEGBEGLVPKYLWRSQ--DVENCAVRGLTDECHYIVRL 128
 QY 125 LKKNDALFVCGTNAFNPSCRNKKMTLEPPGFSPGNARCFYDANKANVALFADGLYS 184
 DB 129 VPMDSQTLACGTNSFSPVCSYGTISLQCEBELSGARCPFDATQGNVALFARGSLYS 188
 QY 185 ATVTDFLAIDAVYSLGESPTLRTVFKHDSKWLKEPPIVQAVDYGDIYFFFRRLAVYVN 244
 DB 189 ATAADFQASDAVAVYSLGPOPPLPSAKYDSKWLREPHFOALHEHDHYVFFRREVSVDA 248
 QY 245 TMGKVVFRVAVOVCKNDMGSGORVLEKQMTSLKRLKNSVGDHFFENILQATVDIR 304
 DB 249 RLGKVFQFRVAVKCRDMGSPRALDRHTSLFKRLKNSVGDSTFFDVQLALGTVN 308
 QY 305 INGRDVNATFSTPNASIPGSAVCAVMDLIDAVTGRFEKQKSPDSTWTPVDERVXP 364
 DB 309 LHGRALRGVFTTQNSIPGSAVCAFYDEIRGEKFKERSLDGMATPVEDRVPSP 368
 QY 365 RFGCCAGSSSLERVATSNFPPDNTNFKTPELMDAVPSIFNRPWFRTWRYRLTKIA 424
 DB 369 RFGSCAGVGGAALFSSSRDLPDVLTFFIKHFLDPAVPVTHQF-LITLSRALLTQVA 427
 QY 425 VTTAAGPYNHTVFLGSEKGIILKFLARIGNSGFLNSLFLFEMSVNSEKSDVGED 484
 DB 428 VGMGPHSNITVMFLGSDGTVLKVLPRGSGG-PEITLLEEDANSFARCSGKRTQ 486
 QY 485 --KRIMGMOIDRASSSLVYAFSTCVIYVPLGRCEHCKCKCTIARSPYCGMIKEGAC 542
 DB 487 TARKITIGLELDTEGRLFLVAFSGCIYVPLSCAHGAGQRCCLASQDPYCGMHSRPG-C 545
 QY 543 SHLSPNSRLTFEODIERGNTDGL--GDCHNSFVALNGHSSLLPSTTSSDSTAQEGYER 600
 DB 546 VDIRSGGTDVQ--AGNQSMEHEDCODG--ATGSGSGPDSAY-- 586
 QY 601 GSKLMDKHLDDPDSSTPLGAVSSHHQKKGVITRESYKGDOLPVVLLAIAVLAIV 660
 DB 587 -----GVRDLPPASASRSVPITLLASVAABA 615
 QY 661 MGAVERGITVYCVCDHRRKDVAVVQKREKELTHSRG-----SMGVTXLSGL 708
 DB 616 LGASVGLLVSCG--RR-----AARRGKOIEFTGLPRPILSRSLARHG- 659
 QY 709 FGDTSQKDPF--FALLTEPMANGKATRGNTAKMLIKADQHHLDLTALPTSESTTLQO 766

DB 660 -GGPEPPSKDGDVQTPOLYTTFLPPEPGVPP-----ELACLPTESTPELV 709
 QY 767 KRRPSRGRWERNONULINACTKMDPMGSPVLPDLPRLASPSH-----IPSVVLLPIT 821
 DB 710 KHLRAAD-PWENQNNMNA-----KEGPGSRGGAAGAPRVLVPR-- 752
 QY 822 QCGYQHEVDQPKASEVAQMALEDOATLEVKTIKEHL---SKSPNHNVLVENLDS- 876
 DB 753 -----PPGCGPQ-----AVETTLLEELRYLHGQPPKGAEPAPLTSR 793
 QY 877 -LPPKVPQREASIGPPASLSQTLGSKYLEMHHSSYGVYDKRSYPTNSLTRSHQATTLK 935
 DB 794 ALPPE--PALLOGSPRHECHSPRLDY-----PEGRGASAPV----- 833
 QY 936 RNNNTSSNSHL-----SRNOSFGRGDNPPAPQVDSIQVHSSQPSGQAVTVSPQS- 988
 DB 834 -RPALSAAPARLVGVGGGRRLPFGSGHAPRALPALTIV-----PSGGPRSYSGPGKHL 883
 QY 989 -LNAVNSLTRSGIKRTPSLKPDVPPKPSF-APLSTSMKPN 1026
 DB 884 LYLGRPEGYRGRLKRVDEKPOLSLKPLVGVSSROAVPN 924

RESULT 7
 SM6C_MOUSE STANDARD; PRT; 931 AA.
 ID SM6C_MOUSE
 AC Q9WNT3;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
 GN SEMA6C OR SEMA Y.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 CX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=99160821; PubMed=10049528;
 RA Kikuchi K., Chedotal A., Hanafusa H., Ujimasa Y., de Castro F.,
 RA Goodman C.S., Kimura T.;
 RT "Cloning and characterization of a novel class VI semaphorin,
 semaphorin Y.";
 RL Mol. Cell. Neurosci. 13:9-23 (1999).
 CC -1- FUNCTION: MAY BE A STOP SIGNAL FOR THE DORSAL ROOT GANGLION
 CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER
 CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF
 CC NEURONAL CONNECTIONS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB013729; BAA76294.1; --
 CC MGD: MGI:1338032; Sema6c.
 CC InterPro: IPR001627; Sema.
 CC Pfam: PF01403; Sema; 1.
 CC SMART: SM00630; Sema; 1.
 CC Signal, Transmembrane, Multigene family; Neurogenesis; Glycoprotein;
 KM developmental protein.
 FT SIGNAL 1 25
 FT CHAIN 26 931 SEMAPHORIN 6C.
 FT DOMAIN 26 605 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 606 626 POTENTIAL.
 FT DOMAIN 627 931 CYTOPLASMIC (POTENTIAL).

```

FT DOMAIN 234 541 SEMA.
FT DOMAIN 663 669 POLY-PRO.
FT DOMAIN 754 757 POLY-PRO.
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 931 AA; 99537 MM; B0D99D54209F125 CRC64;

Query Match 27.4%; Score 1493.5; DB 1; Length 931;
Best Local Similarity 35.7%; Pred. No. 2e-88;
Matches 374; Conservative 145; Mismatches 353; Indels 177; Gaps 32;

QY 6 LLLYFTLLHAGGPEDESEFISISGNVTKQYPVVGHGEGNTQRRHLDIOMIMN 65
DB 13 LLLLSLSPQAALAFQDDPTPLTSDIQAGSPSSWRLGDDVAAL-GLDQRLTLTA 71
QY 66 GTLYIARHDIVYVDIDTSHT-EEIYCSKLLTWKSRQADVDTCRMKGNKDECHNEIKYL 124
DB 72 RTLLVARDHVFSFDIQAGEEGGLVNNKFLTRSG--DWENCAVAGKLTDECYNTIRVL 129
QY 125 LKKNDDALFYCGTNAENPSCRYNKMOTLEPFGEFSGMACPYDAKXANVALFADGKLYS 184
DB 130 VPMNSQTLACGTNSSPMCRSYGITSLOEGEELSGQACPEDATQSTVALFAESGLYS 189
QY 185 ATVTDLAIDAVIYRSIGSEPTLRVYKHSKWLKEPFYQAVDYGVYFPEFEIAYENX 244
DB 190 ATADPOASDAVYRSLGPOPLRSKAYDSKWLREPHFYALHGEHYVFFREVEDA 249
QY 245 TMGVVFPVRAOYCKNDMGSGOVLKEMQNTSPFKARINCSVPDSHFYENIQAVTDIR 304
DB 250 RLGRVCSRVARYCKRDMGSPRALDHWMTSLKRLINCSVPDSHFYFVDTLSLTGPV 309
QY 305 INGRDVLAFFSPYNSIPSSAICADMDLIAVFTGRFEECKSPDSTWPPVDEVPKR 364
DB 310 LHGSALFGVFTTQNTSIPSAVCAFYLDIEGFEKFEKRESLDGAWTPVEDKVPSP 369
QY 365 RPGCGSSSLERYATNEFPDPTLNFKCHPLMDEAVPSIFRPMFLRTMAYRRLTKIA 424
DB 370 RPSGCAVGAASFSQDLPDVLFLFKHPLIDRVPATQPL-LLTTSALTLQVA 428
QY 425 VDTAAGPYQNTVTVFLGSEKGIILKFLARIGNSGFLNLSFLEMSVYNSKCSYDQVED 484
DB 429 VDGAGPHRNTTVFLGSDNGTVLKVLP-CCGSLGSEPIVLEIDAYSHARCS--GKRS 485
QY 485 ----KRMGQDLBRASSLYVAFSTCYIKYPLGRCEHNGCKCTCLASRPYCGWIKEG 540
DB 486 PRAARITIGELDEGHRLEFVAFPGCIVYLSRCARHGAQCRCLASLDPPYCGMHSRG 545
QY 541 ACSHLSPNSRLTEFQDIERGNTDGLGPHCHNSFYALNHSLSLFTSTSTAGYE-S 599
DB 546 CMSIRGP-----GGTD-----VDLTGNS-----TEHGDDQDAGTGS 578
QY 600 RGGMLDMKHLIDSPDSTDLGAVSSHNHODKGVIRSYLKGHDQVPTLLATAVILAF 659
DB 579 QSGGDSAY-----GVRDLSPASASRSIPILLACVAALF 615
QY 660 VMGAVSEGIYVYCVCHRRKDVAVVQKKEKELTSHRGSSSVYKLSGLGDPQSKPKP 719
DB 616 ALGASVSGGLVSCAC--RRAN-----RRRSKDIEPTGJPRPLSLSLARLGGGPEPPPP 669
QY 720 ---EAILTPLMHNKCLATPGTAKMLKADQHNLDLALPPESTPTLQCKRKRSGRSRE 776
DB 670 KDGDAQTPQLYTFLPLPPDGSGSP-----ELACLPPETTPPLPYKHLRASGG-P 719
QY 777 WERNQNLINACT-XDMPMGSPVPIPTDLF--LRASPHIIPSVVVLPTQO--GYOH-- 827
DB 720 WENQNGNNASEGRPRGCGSGAGRAPRVIVAPPBGCGQAVEVTTLELLRLYLHGP 779
QY 828 -----EYVQPKMSEVAQMALEQOATLEKTKIKELLSKSPHGNVLENTL 874
DB 780 QPFRKGEPLASAPFTSRPPASEGASLFDV-----SSMPMDGVPPL-RL 824
QY 875 DSLPPK-----VPQREASIGPPGASLQGTGLSKRLMHSSSYGVYDKRSYPTNSILTRSHQ 930

```

```

DB 825 D-VPPEGRAPASGRPALAPAPRLG-VGGSRL-----PFPT-----HR 862
QY 931 ATTLKRNNTNSNSHLSNRNSFGGNDPPAPQVDS--IOWHSSQBGCAVTSRQS 988
DB 863 A-----PFGILTRVPSGGAPARYSGGPRHLLYLGR-PE 894
QY 989 LNANSLTRSGLKRTPSLKPDV--PPKPS 1015
DB 895 GHRGSLKRVKDVKSLSPKPPASPPQRA 923

RESULT 8
SMC_RAT STANDARD; PRT; 960 AA.
AC Q9WTU3; Q9WTU6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).
GN SEMA6C OR SEMAY.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SOURCE FROM N.A. (ISOPFORM SEMA Y-L AND SEMA Y-S).
RC STRAIN=Sprague-Dawley; TISSUE=Muscle;
RA MEDLINE=99160821; PubMed=10049528;
RA Kikuchi K., Chedotal A., Hanafusa H., Ujimasa Y., de Castro F.,
  Goodman C.S., Kimura T.;
RT "Cloning and characterization of a novel class VI semaphorin,
  RT semaphorin Y.";
RL Mol. Cell. Neurosci. 13:9-23 (1999).
CC - FUNCTION: SHOWS GROWTH CONE COLLAPSING ACTIVITY ON DORSAL ROOT
  CC GANGLION (DRG) NEURONS IN VITRO. MAY BE A STOP SIGNAL FOR THE DRG
  CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER
  CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF
  CC NEURONAL CONNECTIONS.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - ALTERNATIVE PRODUCTS:
  CC Event=Alternative splicing; Named isoforms=2;
  CC Name=Sema Y-L;
  CC IsoId=Q9WTL3-1; Sequence=Displayed;
  CC Name=Sema Y-S;
  CC IsoId=Q9WTL3-2; Sequence=VSP_006048;
CC - TISSUE SPECIFICITY: EXPRESSED IN MANY REGIONS OF THE DEVELOPING
  CC NERVOUS SYSTEM, PROBABLY AS IMMATURE MUSCLE AND DERMIS. IN ADULT,
  CC IN NONNEURAL TISSUE SUCH AS IMMATURE MUSCLE AND DERMIS. IN ADULT,
  CC STRONG EXPRESSION IN THE SKELETAL MUSCLE AND MODERATE EXPRESSION
  CC IN THE BRAIN, WHERE CEREBELLUM SHOWS THE HIGHEST EXPRESSION. ALSO
  CC EXPRESSED IN ALMOST ALL AREAS OF THE CNS.
CC - DEVELOPMENTAL STAGE: DETECTED AT E12 AND FOUND AT MARKEDLY
  CC INCREASED LEVELS AT E15 AND E18 IN BOTH THE HEAD AND THE BODY. AT
  CC BIRTH THE LEVEL DECREASES SIGNIFICANTLY.
CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC - SIMILARITY: Contains 1 Sema domain.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation
  CC at the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb.ch/announce/
  CC or send an email to license@isb.sib.ch).
CC -
DR EMBL; AB000817; BAA76293.2; -
DR EMBL; AB014074; BAA76295.1; -
DR InterPro; IPR001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00630; Sema; 1.
KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
  KW Developmental protein; Alternative splicing.

```

FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 960 SEMAPHORIN 6C.
 FT DOMAIN 24 635 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 636 656 POTENTIAL.
 FT DOMAIN 657 960 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 234 541 SEMA.
 FT DOMAIN 693 699 POLY-PRO.
 FT DOMAIN 783 786 POLY-PRO.
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 586 617 Missing (in isoform Sema Y-S).
 FT /FtId=VSP 006048.
 SQ SEQUENCE 960 AA; 102610 MW; C882933C607E6086 CRC64;

Query Match 27.4%; Score 1491; DB 1; Length 960;
 Best Local Similarity 35.1%; Pred. No. 3,1e-88;
 Matches 374; Conservative 149; Mismatches 363; Indels 180; Gaps 32;

QY 6 LLLVYTLHPAGAPDESEPISSHGNYTKQYVYVGHKFGKRNTPQHRRLDIQIMIMN 65
 DB 11 LLLLLLSPQOTAPFPDPIRLTSDLOGTSSMFRLEDDAAVAEL-GIDPQRFLLIN 69
 QY 66 GTLYIARDHIYVDIDTSHT-EEIYCSKSLTWKSRQADVDTGRMKGKHDECHNFIKVL 124
 DB 70 RTLLVAARDHVSFDLQAESEGLVPRKFLTWRSQ--DMENCAVRGKLTDECYVIRVL 127
 QY 125 LKANDALFVCGTNAFNSCRKYNKDTLEPPGDESSGMAKPCYDAKANVALFADGKLYS 184
 DB 128 VPMDSQTLACGTNFSFVCRSYGATSLQEGEELISGQRCFDTQSTVAISAGSLYS 187
 QY 165 AVTPEFLDAIVAYSLSGESPTLRVTKHDSKWLKEPYEQAQADYDIYFFPREIAVEYN 244
 DB 168 ATADPQASDAVAVYSLGQPLRLSAKYSKWLREPHFYALHEHGDHYFFPREISVDA 247
 QY 245 TWGKVFPRVAQVCKNDMGSGRVLEKQWTSFLKARLNCVSDSHFFENIIQAVTVIR 304
 DB 248 RIGRQFQSVARVACRDMGSGRALDRHMTSFLKRLNCVSDSTFYEDVLQSLTPVN 307
 QY 305 INGRDVVATESTPINSITGSAVCAVMDLIDSVTGRKREKQSDSTWTPVPRDRPKP 364
 DB 308 LHGRSLFGVFTTQNTSIPGSAVCAFYLDLIRGEFGKREKQSLDGAATPVSEDKVBP 367
 QY 365 RRGCGAGSSSLERVATNSFEPPDTNFIKTHPLMDEAVSINRPFELTWTVRYLTKIA 424
 DB 368 RRGSGGAGAAALFSSQDLPPDVLFIQAHPLDPAVPAHQP-LLTLSRALLTYVA 426
 QY 485 ---KRMGQDRASSSLYVAFSTVIAKPLGRCHRGKCKKTCIASDPPCGWKIEG 540
 DB 484 PRAARIIQLEHDTBEGHRLFVAFPGCIYVLSRCRHAQCCORSCIASIDPYCGMHRFG 543
 QY 541 ACHSLSPNRILTEQDIE-RGNTDG--GDCHNSFVALN---GHSSSLP-----PSITT 588
 DB 544 CVNIRPBG---TDVLTGNQESHEHGDQDGAIGSGSGPDSAYVLLGPSPETPS 598
 QY 589 SDSTAOEYVESRGMLDMKHLIDSPDTPLAGVSSHNDKKGVITRESYLKGDQLVPY 648
 DB 599 SPEDAIRPGQS-----STLGA-----HTGGVARDLSPASRSRIP1 634
 QY 649 TLLAIIVILLAFVGVANFSGITVYCVCDHRKDVAVVQKREKELTHRRGMSVTLGSG 708
 DB 635 PLLIACVAAPFALGASVGLVSCAC--RRAN---RRRSKDIETDGLPPPLRLSLARI 688
 QY 709 FGPTGSDPKP---EALITPLMNGKLTAPGTAKMLIKADQHLLDALTALPTESTPTIQ 765
 DB 689 HGGGPPPPPPKGGDAQTQLYTTPRPPGSGSP-----ELACLPTEETTPELP 739
 QY 766 QKRPKSGRSEWRNQLINACTKMPKSGPVIPTDLP---LRASDPLPSVVLPTIO 822

DB 740 VKELRASGG-PWEMNQNNASBGRPRRGCSAAGPAPRVLRPPPGCGQGEVEYTTL 798
 QY 823 Q---GYQH-----EYVDQPMSEVAQALDQAATLEYKTIKEHLSSKSP 864
 DB 799 EELLRYLHGPPQPRKSGEPPLASAPFTRPPASEGALFYD-----SSFMP 844
 QY 865 NRGVNLVENIDSLP-----KVPRBASLGPAGSLSTQGLSKLEMHSSSYG 913
 DB 845 R-----DCEPFLRLDVPDPGKRAAPSGRPLSAPAPRLGVSG--SRRL----- 885
 QY 914 VDYKRSYPTNSILRSHQATTLKNNNTSSNSHLSRQSGFGDPPAPQPVDS--IQV 971
 DB 886 ---PPT-----PFA-----PFGILTIVPSGGGSR 907

RESULT 9

ID SMA_SCHAM STANDARD; PRT; 730 AA.

AC Q26473;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 1A precursor (Semaphorin-I) (Sema I) (Fasciclin IV).
 CN SEMA-1A OR FAS4.
 OS Schistocerca americana (American grasshopper).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyngota;
 OC Neoptera; Orthopteroidea; Orthoptera; Cellifera; Acridomorpha;
 OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
 OX NCBI_Taxid=7009;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93040225; PubMed=1418998;
 RA Kolodkin A.L., Nathans D.J., O'Connor T.P., Patel N.H., Admon A.,
 RA Bentley D., Goodman C.S.;
 RT "Fasciclin IV: sequence, expression, and function during growth cone
 RT guidance in the grasshopper embryo";
 RT Neuron 9:831-845(1992).
 CC - FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - TISSUE SPECIFICITY: DYNAMICALLY EXPRESSED ON A SUBSET OF AXON
 CC PATHWAYS IN THE DEVELOPING CNS AND ON CIRCONFERENCE BANDS OF
 CC EPITHELIAL CELLS IN DEVELOPING LIMB BUDS.
 CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC - SIMILARITY: Contains 1 Sema domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL, L00709; AAA29808.1; -;
 CC PIR: JH0798; JH0798.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF01437; PSI; 1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR Signal; Developmental protein; Transmembrane; Glycoprotein;
 KW Neurogenesis.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 730 SEMAPHORIN 1A.
 FT DOMAIN 21 630 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 631 651 POTENTIAL.
 FT DOMAIN 652 730 CYTOPLASMIC (POTENTIAL).

```

FT DOMAIN 215 515 SEMA.
FT DOMAIN 217 222 POLY-RHE.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 730 AA: 81214 MW: 66655946E7F8E57F CRC64;

Query Match 17.8%; Score 971.5; DB 1; Length 730;
Best Local Similarity 34.4%; Pred. No. 5,9e-55;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

8 LFTLLHFAAGFPDESEPISSHGNYTQYPFVGHKGRNTQK-----HRLDIQMI 61
11 LFWVAH--AAAVNDVSP-----KMYVOF-----GEERVOFLNGSHKHDFKLL 54
62 MINNGTLYIAARDHIYVDID--TSHTTEIYCSKLTWRSQADVTCMKGKHDECHN 119
55 EKDNHSLIVGARIVYNISLRDLTFTEQ-----RIEMHSSGHRRLCYLKGKSEDDCN 109
120 FIKVLLKNDLALFVGTNAFNSCKNYKMDLEPGD-----EFGMARCPIDAHNAV 174
110 YIRVLAKIDDDRVLLICGTNAKPLCRHYALKD-----GDYVVEKEVGRGLCPDPDHNST 165
175 ALFADGLVATVDTFLAIDAVYRSLGSEPTLRVTKHDSKMLKEPYVOAVDGYIYF 234
166 AIYSEGLISATVADSSGNDPLIYRG-----PLRTERSDUKOLNANFNVTMYNDFIIF 220
235 FFERIAVEYNTWKRVVPRVAQVCKNDMGSGQVLEKQWTSFLKARLNSVPGDSHFYRN 294
221 FFERIAVEYNTWKRVVPRVAQVCKNDMGSGQVLEKQWTSFLKARLNSVPGDSHFYRN 279
295 ILQAVDVIRIN-GRDV---VLATFSTPYNSIGSVAACADMILASVFTGRKEDKSPD 350
280 EIQSTSDIIEGNYGGQVEKLIYGVFTTPVNSIGSVAACAFSMKSIIESFPGPEKDETMN 339
351 STTPVDEDERVPRPCGCCAGSSSLERVATSNFEPDITLFIKTHPLMDEAVSIFNRPV 410
340 SNMLAVSLKVPERRPQCQND-----SRLPLDVSVNFVSHLMDAVALFAFTFRL 391
411 FLRTMVEYRLTKIADV---TAAQPYQNTVFLGSEKGIILFLARIGNSGFLN----- 461
392 LIRSLQYRFTKLIADVQVARTPDG--KAYDVLFIQTDGKVIKAL-----NSASFSDSDYV 445
462 DSLFLEMYSYNSKESGVDGVEDKRTKMGQDLDAASSLTYAFSTCVIKVYLGRC--ERRG 519
446 DSVVIEELQVLP-----PGVPVKNLYVVRMDGDSKLVVVSODELITAIKLRGCSDKIT 499
520 KCKKTCTIARSDPYCGWIKEGGAGSHL-SPN---SRLTFPQDIERGNTDGLGDCNHFVAL 575
500 NCRB-CVSLQDPYCAMNVELKCTAVSPMSACKRFIONISLGEHKACGGRPQETIV- 557
576 NGRSSSLLPSTTSSTDA-----QEGYSRGGMIDMKLLDPSDSTDLPLAVASHNQ 628
558 -----ASVPYTOPPTKSSGDPVSHIGAFEPPE--IDNEIIVIGVDDSNVLPNTLAEINHA 610
629 DKGVIVESTYLGKHDQVLPV---TL-LAI-----AVILAFVWGANES 666
611 GSK-----LPSSQEKPLPYTAFTLTITAVTSCLAGLVVGFISGFLFS 652

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94094332; PubMed=8269517;
RA Kolodkin A.L., Matthews D.J., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
growth cone guidance molecules."
RN Cell 75:1389-1399(1993).
RP SEQUENCE FROM N.A.
RX MEDLINE=20156006; PubMed=10731132;
RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA Geotige R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Gerhart W.M., Glasser K.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Giodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.D., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mikhlov G., Mislina N.V., Mobarry C., Morris J., Mostrelet A.,
RA Mount S.M., Moy M., Murphy L., Murthy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
RA Spier E.G., Spradling A.C., Stapleton W., Strong R., Sun E.,
RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Maesman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye U., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.C.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
CC -1- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED BY SUBSETS OF NEURONS AND MUSCLES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AROUND STAGE 10, PRIMARILY
CC IN THE DEVELOPING CNS. IN STAGE 16 EMBRYOS, IT IS EXPRESSED AT
CC HIGHEST LEVELS THROUGHOUT THE CNS, AND WEAK EXPRESSION IS SEEN IN
CC PORTIONS OF THE PERIPHERAL NERVOUS SYSTEM, MOST CLEARLY IN THE
CC LATERAL SENSORY CLUSTERS.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L26082; AAA88789.1; -

```


FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 772 AA, 88808 MW, 240907812FEF2D2 CRC64;

Query Match 15.7%; Score 858; DB 1; Length 772;

Best Local Similarity 32.3%; Pred. No. 1,3e-47;
 Matches 209; Conservative 104; Mismatches 251; Indels 84; Gaps 20;

QY 47 GRNTTQHRIDIMIMNGT-----LYIARDHIYVYDIDT 83
 Db 25 GKNVNPRLKLYKEMLENNVITFENGLANSSVHTLDEBRRLVYAGADHIFSNL-- 82
 QY 84 SHREIYCSKLTWKSKQADVDTCRMKGR-KDECHNFTVLLKDDALFVCGTAPNP 142
 Db 83 ---VNKDFGKIYWPVSYTRDECKWAGKQILKECNFTKVLANYQTHLYAGGTAFHP 139
 QY 143 SCR-----NYKMDTLEPFGEF--SGMRCFYDAKANVALFADGKLYSATVDTFAIDA 195
 Db 140 ICTYIEVGHHPEDNIFKLDQSHENGKSPYDPKLLTASLIDGELVSGTADPMGRDP 199
 QY 196 VIYRSLGESEPTLTATVKDSKMLEPYFVQA-----VDYGYIYFPRRIAYENTMGK 248
 Db 200 AIFRTIGHHPRTQCHDSRWLNDPPIAHLLPESDNPEDDKYVEFFRENALDGHSGK 259
 QY 249 VVEPRVAQYCKNDMGSGQVLEKQWTSFLKARLNCVPG---DSHFYFNIQAATVDVR 304
 Db 260 ATHARIGQICKNPFPG-HRSLVKNWTFELFARICSPGNGIDTHP---DELQVFL 313
 QY 305 INGRD-----VLATFTPYNSIPGSAVCAYMDLISVFTGRKEQKSPDSTWTPVDER 360
 Db 314 MNSKDPKNPIYGVFTTSSNIFGSAVCYMSDVAVRFGLPYAHRDGPYQWVPY--QGR 372
 QY 361 VVPRPGCCGSSSLERVATNSPEPDTLNFITHPMDAVSIFPRPELFTMYRKL 420
 Db 373 VVPRPGTGP-SKTFGGFSTKDLFDVITFASHPMYVPPINRPMITDYNQY 431
 QY 421 TKIADVDAAPYQNHVTFVFLGSEKGIILKFLARIGNSGFLDLSFLEMSVNSSEKSYD 480
 Db 432 TQIVDVRVADGQDYMEIGTGVYLVKAVSPKEMHLEVLLEMMVFR----- 484
 QY 481 GVEDKRIWQMDLRASSSLVYAFSTCVIKVPLGRCERHGGCKTCTIARDPYCGWKIEGG 540
 Db 485 --PPTISAMELSTKQOQLYIGSTAGVAPLHRCDIYGACABCCCLARDPYCAM--DGS 540
 QY 541 ACSHLSNS-RLTFEQIERNNDGLGDC-----HNSFVALNHS--SSLPEPTTSDST 592
 Db 541 SCSRYFPTARRRTRRQDIRNG--DPLTHCSDLQHHNH--HGSLERILYGVENSSTP 595
 QY 593 ADEGESRGMLDMKHLSDPSTDPPLGAVSHNHODKKVIRESYLK 640
 Db 596 LECGPKQKRALVWQFORNEDRKEEL-RVGDHILIRTEQGLLRLSLQK 642

RESULT 12
 SM3A MOUSE STANDARD; PRT; 772 AA.
 ID O08655; Q62180; Q62215;
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D) (Sema D).
 GN SEMA3A OR SEMAD OR SEMD.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI; TISSUE=Embryo;
 RX MEDLINE=95267431; PubMed=7748561;
 RA Pleschel A.W., Adams R.H., Betz H.;

RT "Murtine semaphorin D/collapsin is a member of a diverse gene family
 RT and creates domains inhibitory for axonal extension.";
 RT Neuron 14:941-948(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97470885; PubMed=9311345;
 RA Tanguuchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
 RT Yagi T.;
 RT "Disruption of semaphorin III/D gene causes severe abnormality in
 RT peripheral nerve projection.";
 RT Neuron 19:519-530(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kimura T., Fishman M.C.;
 RT "cDNA sequence of mouse collapsin/semaphorin III.";
 RT Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 107-772 FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=95267432; PubMed=7748562;
 RA Messersmith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,
 RT Goodman C.S., Kolodkin A.L.;
 RT "Semaphorin III can function as a selective chemorepellent to pattern
 RT sensory projections in the spinal cord.";
 RT Neuron 14:949-959(1995).
 CC -1- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO
 CC PATTERN SENSORY PROJECTIONS BY SELECTIVELY REPELLING AXONS THAT
 CC NORMALLY TERMINATE DORSALLY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT
 CC (E11) IN DISTINCT REGIONS OF THE NEUROECTODERM AND MESODERM.
 CC EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.
 CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X85983; CAA59985.1; -
 CC EMBL: D85028; BAA19773.1; -
 CC EMBL: LA1541; AAT77611.1; -
 CC EMBL: LA0484; AAA73934.1; -
 CC PIR: I48747; I48747.
 CC PIR: I58169; I58169.
 CC MGI: MGI:107558; Sema3a.
 DR InterPro: IPR007110; Ig-like.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR001627; Sema.
 DR Pfam: PF00947; Ig_1.
 DR Pfam: PF01403; Sema; 1.
 DR SMART: SM00409; Ig_1.
 DR SMART: SM00423; PSI; 1.
 DR SMART: SM00630; Sema; 1.
 DR PROSITE: PS00835; Ig_Like; 1.
 DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
 KM Developmental protein; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 772
 FT DOMAIN 240 538
 FT DOMAIN 579 665
 FT DOMAIN 728 770
 FT DISUFID 650 723
 FT CARBOHYD 53 53
 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 193 193 D -> N (IN REF. 4).
FT CONFLICT 207 207 H -> D (IN REF. 1).
FT CONFLICT 253 253 D -> G (IN REF. 1).
FT CONFLICT 352 352 F -> L (IN REF. 4).
FT CONFLICT 403 403 A -> G (IN REF. 1).
FT CONFLICT 571 572 OH -> ED (IN REF. 1).
FT CONFLICT 616 620 EDKKE -> RRSKR (IN REF. 1).
FT CONFLICT 623 623 R -> K (IN REF. 4).
SQ SEQUENCE 772 AA, 88799 MW, E89A08528B10AEC3 CRG64;

Query Match 15.7%; Score 853; DB 1; Length 772;
Best Local Similarity 32.2%; Pred. No. 2.7e-47;
Matches 208; Conservative 102; Mismatches 257; Indels 78; Gaps 18;

QY 47 GRNTQRHRLDQIMINMG-----LYIARAHLYVDIT 83
DB 25 GKNVPRRLKLSYKELSNVITFENGLANSSHYTFLLDEERSRLYVGAKHIFSENL-- 82
QY 84 SHTEIYSCSKLTKMSRQADVTCRMKGX-KDECHNFIKVLKKNDALFVCGTNAFNP 142
DB 83 ---VNIKPFQKLVWVSTYTRBECKMAGKDIKECANFKVLEAVNQHLYACGTGAFHP 139
QY 143 SCR-----NYKMDLLEPFDEF--SGMARCPYDAKHANVALFADGKYSATVDFLAIDA 195
DB 140 ICTYIEVGHHPEDNIFKLQDSHFENGRGKSPYDPLKTLASLIDELXSGTADPMGDF 199
QY 196 VYRSLGSEPTLRKYHDSKMLKEPYVOA-----VDYGYITFFREIAYENVMGK 248
DB 200 ALFRLGHHHPRTQHDHSMRLNDPRFISAHILPESDNEDDKVYFFRENIDGESHGK 259
QY 249 VVEPVAQCVKMDMGSCVLEKQWTSFLKARLNSVPG---DSHFYFNILQAVTVYR 304
DB 260 ATHAIGQICRNDPFG-HRSLVNMKTTPLKARLISGVOPNIDHF-----DELQDVF 313
QY 305 INGRD---VVIATSTPYNSIPSGAVCAVMDLIASVTFGEKQKSPDSTWTFVDPDR 360
DB 314 NMSKDPKPIYGVFTTSSNIFKGSAGVMSVSDVRVFLGPYARHDGPNYQWVY-QGR 372
QY 361 VKEPRPGCCAGSSSLERYATSNBPDDTLNFKTHPLNDEAPSPFNENPWFRTVRYRL 420
DB 373 VYPRPGTGP-SKITGGEFSTDLDPDVTTFARSHAMNPPFPNNRIMIKTVNQF 431
QY 421 TKIADTAAGPYQNTVVFLESEKGIKFLARIGNSGFLNSLFEKSVYNSSEKSYD 480
DB 432 TQIVDVRVADGQYDWFIGTGVVGLKVSVPRMETHWDLDEVLLEMTYVR----- 484
QY 481 GVEDKRIMGMQIDRASSSLVYAFSTCVIRKPLGRCEHSGKCKKTCIASRDPYCGMKGG 540
DB 485 --EPTTISAMELSTKQOQLYISTAGVADQLPHRCDIYGKACAECCCLARDPYCAW--DGS 540
QY 541 ACSHLSFNS-RLTFEDIDIRKNTDGLGDCHNSFVALNGSSSL---LPSTTSSTAGE 595
DB 541 SCSSRYFPYAKRTRKODIRNG--DPLTHOSDIQHDHNGHGFLEERIIVGENSSFTLEC 598
QY 596 GYESRGMDMKHLIDSPDTPLAGVSSHNDQKGVIRESYLK 640
DB 599 SPKSGRALVYVQFORNEDRKEFI-RMGDIIRTEGGLLRSLQK 642

RESULT 13
SMA3A HUMAN STANDARD; PRT; 771 AA.
AC 014563:
DT 30-MAY-2000 (rel. 39, Created)
DT 30-MAY-2000 (rel. 39, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
GN SEMA3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=94094332; PubMed=8269517;
RA Kolodkin A.L., Matthes D.J., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
RL growth cone guidance molecules."
RN Cell 75:1389-1399 (1993).
RP SEQUENCE OF 1-37 FROM N.A.
RA Moesner J., Mink P., Hinds K., Strommatt C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 39-182 FROM N.A.
RA Ruffing T., Tin-wollan A.M., Duckels G.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
CC NEUROFILIN-1/PLEXIN-1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: secreted (By similarity).
CC -1- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL, L26081, AAA65938.1, -.
DR EMBL, AC004451, -. NOT ANNOTATED. CDS.
DR EMBL, AC004848, AAC78622.1, -.
DR PIR, D49423, D49423.
DR Genew: HGNC:10723; SEMA3A.
DR MIM, 603961;
DR GO, GO:0005576; C:extracellular; TAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam, PF00047; Ig_1.
DR Pfam, PF01403; Sema_1.
DR SMART, SM00409; IG_1.
DR SMART, SM00423; PSI_1.
DR SMART, SM00630; Sema_1.
DR PROSITE, PS50835; IG-LIKE; 1.
DR Signal, Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 771 SEMAPHORIN 3A.
FT DOMAIN 580 664 SEMA.
FT DOMAIN 727 769 ARG/XYR-RICH (BASIC).
FT DISULFID 649 722
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 771 AA, 88889 MW, 9985F8D3E8ED8456 CRG64;

Query Match 15.6%; Score 852.5; DB 1; Length 771;
Best Local Similarity 32.1%; Pred. No. 2.9e-47;
Matches 208; Conservative 106; Mismatches 250; Indels 83; Gaps 20;

QY 44 HKPGNTQRHRLDQIMINMG-----LYIARAHLYVD 80

```

```

Db 22 YONGKUNVPRKLKLYKEMLENNVTPFENGLANSSSYHTFLDEBSRLVYGAKHIFSPD 81
Qy 81 IDTSHTBEIYCSKLTWKSQADVDTCRMKGX- KDCBHFIFKVLKKNDDALFVCGTNA 139
Db 82 L-----VNIDFQKITWPSVSYTRDECKMAGKQILKECAFIVLKVAYNTHLYAGGTG 136
Qy 140 FNPSCR-----NVMQDTLEPFGEDEF--SGMARCPYDAKANVLAFLADGKLYSATVDFLA 192
Db 137 FHPICLYIEIGHNPEDNIFLENSHPENGSGSPYDKLTLAALLIDGELYSSTAADFNG 196
Qy 193 IDAVIYKSLGESPTLRITVKHDSKWLKEPYVQA-----VDYGDYIYFFPRELAVEYNT 245
Db 197 RDAIFRTGLHHHPIRTEGDSRLNDPKFISALLISEDNPPDDKYFFEFENALDGEH 256
Qy 246 MGKVPFPRVAVOCCKNDMGSSQVLEKQWTFKLARLNGSPG-----DSHYFNILQAVTD 301
Db 257 SGKATHARIQIQIKNDPFGG--HRSIVNMKTFELKARLISGVPNGGIDTFR-----DELD 310
Qy 302 VIRINGRD-----VLATFSTPYNISPGSAVCAYDMLDIASVFTGRFKEQKSPDSTWTPVP 357
Db 311 VFILNFKDPKPNVYVYVFTTSSNIFKGSAYCWTSMVSRVFLGPRYAHKDPYQWVY- 369
Qy 358 DERVPKRPCCAGSSSLERVATSNPPDDTLNFIKTHPLMDEAVPSIFNRPWFLRTWYR 417
Db 370 QGRVPYPRPGTGP--SKTFGGFSDSTKDLDDVITFARSHPMYVNPVFPNNRPVYIKTDVN 428
Qy 418 YRLTKIADVTAAGRYQNHVTFELGSEGIILKFLARLNGSGFLNDSLFLEEMGVNSEKC 477
Db 429 YQFQIYVDAVDADGQYDWFVIGTDVGVTKVVISIPKETWYDLEELLEMVTFR----- 484
Qy 478 SYDVEDERKIMGNOLDRASSSLVYAFSTCVIKVPLGRCERHKKCKTICIASRDPYCGMIK 537
Db 485 -----EPFAISAMELSTKQOOLYIGSTAGVAGQPLHNCIDYGRKACACCLARPYCAM-- 537
Qy 538 EGAGCSHLSFNS-ELTEQDIERGNTDGLGDC-----HNSFVALNGHS--SSLPSTTSD 590
Db 538 DGSACSRVFPFAKRTBRQDIRNG--DPLTHCSDLHHDNH---HGSPEERLIYGVENSS 592
Qy 591 STAGEVYESRGMTDMKHLSDSPDTPPLGAVSSHNDKKGVIREG 637
Db 593 TFLCSPKSORALVYQFORNERKEEI--RVDHIIIRTQGLLRS 638

RESULT 14
SKLA_TRICE STANDARD; PRT; 712 AA.
ID SKLA_TRICE Q26972;
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 1A precursor (Semaphorin-1).
GN SEMA-1A OR TSEMA-1.
OS Tribolium confusum (Confused flour beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tribolium.
NCBI_Taxid=7071;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94094332; PubMed=8269517;
RA Kolodkin A.L., Mathes D.U., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
RT growth cone guidance molecules."
RL Cell 75:1389-1399(1993).
CC -1- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 Sema domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; L26080; AAA16609.1; -
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
KM Signal; Developmental protein; Transmembrane; Glycoprotein;
KM Neurogenesis. 1
FT SIGNAL 1 20
FT CHAIN 21 712
FT DOMAIN 21 601
FT TRANSMEM 602 622
FT DOMAIN 623 712
FT DOMAIN 213 506
FT DOMAIN 470 473
FT DOMAIN 612 616
FT CARBOHYD 42 42
FT CARBOHYD 69 69
FT CARBOHYD 161 161
FT CARBOHYD 265 265
SQ SEQUENCE 712 AA; 79751 MW; C0734FE5B9C09E93 CRG64;

Query Match 15.5%; Score 845; DB 1; Length 712;
Best Local Similarity 28.9%; Pred. No. 8e-47;
Matches 230; Conservative 130; Mismatches 270; Indels 166; Gaps 30;

12 LILFAGAPPEDESEPT-----SISGNYTKQYVFGHKRGNTQRHRLDIQWIMIMN- 65
12 LIALCHAMPDSSSKLINHPKSVESKST-----GNATFPDH-----FLYLNQ 54
Qy 66 --GLTYIARADHIYTVDI--DTSHTBEIYCSKLTWKSQADVDTCRMKGXKDCBHFIFK 122
Db 55 DETSILVGGNRRYNNISIFPLSERK-----GGRIDWSPDAHGQLCLTKGATDDCCQYIR 110
Qy 123 VLIKKNDDALFVCGTNAFNSGCRNYKMDTEPEPGD--EFGSMACPDYDAKANVLAFLADGK 161
Db 111 ILYSSEPGKLVIGTNSYKPLCRTYAFKQKYLVEKEVEIGICFPYNNHNSSTSVSNGQ 170
Qy 182 LVGATYTDPLAIDAVIYRSLGESPTLRITVKHDSKWLKEPYVQAVDYGVIYFFPREIAV 241
Db 171 LFSATVADBSGGPPLYRE-----PQRTLSIDLKOLNAPFVAVAGDYLFFFYRETA 225
Qy 242 EYNTMGKVPFPRVAVOCCKNDMGSSQVLEKQWTFKLARLNGSVPGDSHYFNILQAVTD 301
Db 226 EYNNCGKVIYSRAVAVCKDKDGGPHQSRDR--WTSFLKARLNGSIPGEYFPYPEIOSTSD 284
Qy 302 VI--RINGRP-----VLATFSTPYNISPGSAVCAYDMLDIASVFTGRFKEQKSPDSTWTPV 356
Db 285 IVEGRVNSDSCSKITIIIGILTPVNAIGSALICVQVADILRVBGSFKHOFETINSMLPV 344
Qy 357 PDERVPKRPCCAGSSSLERVATSNPPDDTLNFIKTHPLMDEAVPSIFNRPWFLRTWYR 416
Db 345 PQLNIVPEPRPGQCYRDSRI-----LPKQNVNFIKTHLMD--VALGKPLVAVSL 395
Qy 417 RYRLTKIADVTAAGRYQNH--TVFELGSEGIILKFLARLNGSGFLNDSLFLEEMSVYNS 474
Db 396 QYRTATITVDPOVKTINNQYLDVLYIGTDGKVLK----- 430
Qy 475 EKCYSDVEDERKIMGNOLDRASSSLVYAFSTCV-----IKVP----- 511
Db 431 -----ANITPRHAKKALVYKRTISVHPHAPYQKLIKLAGYGVKVVYVGKDEI 478
Qy 512 ---LGRCEHGGCKTKCIASRDPYCGMIKEGAGCSHLSFNSRLTF--EQDIERGN----- 561
Db 479 RLALNLNHCASKTRC--KQCVLEQDPHCAMDAKQNLVCVSDIVTSYRFLIQDVVRGDNDKWC 537

```

```

QY 562 ----TDGDCCHNSFVALNGHSSSLPSTTSDSTAQEGYESRGMLDMKHLDSPDSTDP 618
Db 538 SPQTDKKTIVK-----PSEVENET-----NSIDEDCL-----DSSDP 572
QY 619 LGAVSSHNDKQKVIRESYLKG---HDQIVPV---LTAIVILAFVMA---VFEGITV 670
Db 573 LKGTGLDDSDSCPV-SENSIGCAVRQGLVIYTAGTLHIWVWVSVIGLFSMLVSLGSLV 631
QY 671 YCV--CDHRKDVAVVQKE--KELTHSRGNS-SVTLSCGFQDQSKDPKPEALITP 725
Db 632 FAKFHSQYPAFPFLEQNHLERLSANQGYLTPRANKAVNLVVKVSSSTPRPKMD 691
QY 726 LHMNGKATPANTAKM 741
Db 692 VKKDINIASDGLQKI 707

RESULT 15
SZIB BRARE STANDARD; PRT: 778 AA.
AC 09M66;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin ZIB precursor (Semaphorin 1B) (Sema-ZIB).
GN SEMAZIB OR SEMAZIB.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99425174; PubMed=10495275;
RA Roos W.; Schachner W.; Bernhardt R.R.;
RT "Zebrafish semaphorin ZIB inhibits growing motor axons in vivo.";
RL Mech. Dev. 87:103-117(1999).
CC - FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
CC - SUBCELLULAR LOCATION: Secreted (by similarity).
CC - TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
CC VENTRALLY EXTENDING MOTOR AXONS.
CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC - SIMILARITY: Contains 1 Sema domain.
CC - SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: AF083382; AAD28103.1; -
DR ZFIN: ZDB-GENE-991209-6; sema3ab.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003599; IG.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; plexin_repeat.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01437; PSI_1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
DR PROSITE: PS50835; IG-LIKE; 1.
DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.

```

```

FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 778 SEMAPHORIN ZIB.
FT DOMAIN 241 539 SEMA.
FT DOMAIN 579 668 IG-LIKE C2-TYPE.
FT DOMAIN 721 776 ARG/LYS-RICH (BASIC).
FT DISULFID 652 716 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 593 593 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 778 AA; 88904 MW; 4D36FA323AE21895 CRC64;

Query Match 15.5%; Score 845; DB 1; Length 778;
Best Local Similarity 34.3%; Pred. No. 9e-47;
Matches 194; Conservative 94; Mismatches 204; Indels 74; Gaps 17;

QY 66 GTLYIARDHIYTVD-IDTSHTLEYCSKLTWKSQAQADVTCRMKGRH-KDECHNFIKV 123
Db 66 GRLFYGAKDHYLSFNLVDINNDDQL-----ISWPSRPRDECKAKGADVCKECAFIV 122
QY 124 LKKNDDALFVCGTNAFNPSCRNYKM-----DLEPPGDEF-SGMARCPYAKHANVALF 177
Db 123 LQPFNQTHLYACGTGAFHPCAHAVEGKRSDDNFRLLGSSFENGKSPYDPKLOTASVL 182
QY 178 ADGKLYSATVNDPLAIDAVIYRSIGESPLRTLVKHSKWLKEPPYV-----QAVDY 228
Db 183 IDGELYAGTSADFPKGRDPALFRLLGKHPITREQDSRLNDPRFVSVHLIPESDVAED- 241
QY 229 GDYIYFFPREIAVENYNTGKVFPRVAVQVCKNDMGSGRVLEKOWTSFLKALNCSVPG- 287
Db 242 -DKIYLFPRENALIDDEQISKATYHARIGQLCKNDFGG-HRSLVKNKTTFLKALVCSVPL 299
QY 288 ---DSHFYFNILQAVTVDIRNGRD---VVLATSTYNSIPGSAVCAVMDLDIASVPT 340
Db 300 NGIDTHF-----DELQDVFLMSKDPKXNPITYAAFTTSSNIFKGSAGVCMYSMADIRRVFL 354
QY 341 GRFEQKSPDSTWTPVDPDERVYKPRPGCCAGSSSLERYATSNFEPDDTLNFKTHPLMDE 400
Db 355 GPYAHKDPNYPQWPEFLN-RVPIYRPGTC-SKITDGESEKTFEDDVITPARSHIPAMYN 412
QY 401 AVPSIFNEPFWLRTWVRRLTKIADVTAAGYQNHHTVFLGSEKGIILKFLARINSGFL 460
Db 413 PVPPINNHPIIKTDVDYQFOIVVDREABDGYDVAFIGTMDQTVLKVSIIPGTWHD 472
QY 461 NDSLFLEMSYVNSEKCSYDGEVDEKRTMGWLDRASSSLVYAFSTCYIKVPLGRGERGK 520
Db 473 LBEVLEEMTVFR-----BPTALTAMELSTKQGLYGSATGVSQMPULHRCDVYVK 523
QY 521 CKTICIASRDPYCGMIKKGACSHSPNS-RLTFEODIERGNTDGLGDCCHNSFVALNGHS 579
Db 524 ACARCCCLARDPYCAM--DGSQCSRYFPYAKRRTTRQDIRND----- 563
QY 580 SLLPSTTSDSTAQEGYESRGMLD 605
Db 564 ----FLTQCSDLQHHDADGEGAGLDD 585

```

Search completed: October 23, 2003, 17:09:37
Job time : 22 secs